LOAN DOCUMENT

PHOTOGRAPH THIS SHEET

390743 SYSP9

AD-A254 459

DTIC ACCESSION NUMBER

LEVEL

INVENTORY

①

AFOSR-TR-92-0705

DOCUMENT IDENTIFICATION

25 Jun 92

DISTRIBUTION STATEMENT

ACCESSION FOR

NTIS	GRA&I	☑
DTIC	TRAC	☐
UNANNOUNCED		☐
JUSTIFICATION		

BY

DISTRIBUTION/

AVAILABILITY CODES

DISTRIBUTION	AVAILABILITY AND/OR SPECIAL	
A-1		

DISTRIBUTION STAMP

DTIC QUALITY INSPECTED 5

DATE RECEIVED IN DTIC

DTIC
S ELECTE
AUG 3 1992
C D

DATE ACCESSIONED

DATE RETURNED

H
A
N
D
L
E

W
I
T
H

C
A
R
E

92-20727

REGISTERED OR CERTIFIED NUMBER

PHOTOGRAPH THIS SHEET AND RETURN TO DTIC-FDAC

DTIC FORM 70A
JUN 90

DOCUMENT PROCESSING SHEET

PREVIOUS EDITIONS MAY BE USED UNTIL
STOCK IS EXHAUSTED.

LOAN DOCUMENT

390743 5457g

AD-A254 459

UNITED STATES AIR FORCE

1989 RESEARCH INITIATION PROGRAM

Conducted by

UNIVERSAL ENERGY SYSTEMS, INC.

under

USAF Contract Number F49620-88-C-0053

RESEARCH REPORTS

VOLUME II OF IV

Submitted to

Air Force Office of Scientific Research

Bolling Air Force Base

Washington, DC

By

Universal Energy Systems, Inc.

June 1992

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	25 Jun 92	ANNUAL 1 Jan 90 to 31 Dec 90

4. TITLE AND SUBTITLE

US Air Force 1989 Research Initiation Program Conducted by Universal Energy Systems, Inc, VOL # 2

5. FUNDING NUMBERS

F49620-88-C-0053
2305/D5

6. AUTHOR(S)

Rod Darrah

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Universal Energy Systems, Inc
Dayton OH

8. PERFORMING ORGANIZATION REPORT NUMBER

AFOSR 92 0705

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

AFOSR/NI
Bolling AFB DC

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT

(U)

12b. DISTRIBUTION CODE

(U)

13. ABSTRACT (Maximum 200 words)

This program is for follow-on research efforts for the participants in the Summer Faculty Research Program. Funding is provided to establish RIP awards to about half the number of participants in the SFRP. Participants in the 1989 SFRP competed for funding under the 1989 RIP. Evaluation of the proposals were made by the contractor. Evaluation criteria consisted of: 1. Technical excellence of the proposal 2. Continuation of the SFRP effort 3. Cost sharing by the university. The list of proposals selected for award was forwarded to AFOSR for approval of funding and for research efforts to be completed by 31 December 1990. The following summarizes the events for the evaluation of proposals and award of funding under the RIP. A. RIP proposals were submitted to the contractor by 1 November 1990. The proposals were limited to $20,000 plus cost sharing by the universities. The universities were encouraged to cost share, since this is an effort to establish a long term effort between the Air Force and the university. B. Proposals were evaluated on the criteria listed above and the final award approval was given by AFOSR after consultation with the Air Force Laboratories. C. Subcontracts were negotiated with the Universities. There were a total of 122 RIP awards made under the 1989 program.

14. SUBJECT TERMS

15. NUMBER OF PAGES

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
(U)	(U)	(U)	(U)

UNITED STATES AIR FORCE

1989 RESEARCH INITIATION PROGRAM

Conducted by

UNIVERSAL ENERGY SYSTEMS, INC.

under

USAF Contract Number F49620-88-C-0053

RESEARCH REPORTS

VOLUME II OF IV

Submitted to

Air Force Office of Scientific Research

Bolling Air Force Base

Washington, DC

By

Universal Energy Systems, Inc.

June 1992

TABLE OF CONTENTS

INTRODUCTION

Research Initiation Program - 1989

AFOSR has provided funding for follow-on research efforts for the participants in the Summer Faculty Research Program. Initially, this program was conducted by AFOSR and popularly known as the Mini-Grant Program. Since 1983 the program has been conducted by the Summer Faculty Research Program (SFRP) contractor and is now called the Research Initiation Program (RIP). Funding is provided to establish RIP awards to about half the number of participants in the SFRP.

Participants in the 1989 SFRP competed for funding under the 1989 RIP. Participants submitted cost and technical proposals to the contractor by 1 November 1989, following their participation in the 1989 SFRP.

Evaluation of these proposals were made by the contractor. Evaluation criteria consisted of:

1. Technical excellence of the proposal
2. Continuation of the SFRP effort
3. Cost sharing by the university

The list of proposals selected for award was forwarded to AFOSR for approval of funding. Those approved by AFOSR were funded for research efforts to be completed by 31 December 1990.

The following summarizes the events for the evaluation of proposals and award of funding under the RIP.

A. RIP proposals were submitted to the contractor by 1 November 1989. The proposals were limited to $20,000 plus cost sharing by the universities. The universities were encouraged to cost share, since this is an effort to establish a long term effort between the Air Force and the university.

B. Proposals were evaluated on the criteria listed above and the final award approval was given by AFOSR after consultation with the Air Force Laboratories.

C. Subcontracts were negotiated with the universities. The period of performance of the subcontract was between October 1989 and December 1990.

Copies of the final reports are presented in Volumes I through IV of the 1989 Research Initiation Program Report. There were a total of 122 RIP awards made under the 1989 program.

STATISTICS

PROGRAM STATISTICS

Total SFRP Participants	168
Total RIP Proposals submitted by SFRP	132
Total RIP Proposals submitted by GSRP	2
Total RIP Proposals submitted	134
Total RIP's funded to SFRP	94
Total RIP's funded to GSRP	2
Total RIP's funded	96
Total RIP Proposals submitted by HBCU's	9
Total RIP Proposals funded to HBCU's	5

LABORATORY PARTICIPATION

Laboratory	Participants	Submitted	Funded
AAMRL	12	10	6
WRDC/APL	10	8	6
ATL	9	9 (1 GSRP)	9 (1 GSRP)
AEDC	10	8	8
WRDC/AL	7	5	4
ESMC	0	0	0
ESD	3	2	1
ESC	11	8	7
WRDC/FDL	9	7	5
FJSRL	7	5	4
AFGL	12	10	6
HRL	12	10 (1 GSRP)	8 (1 GSRP)
WRDC/ML	9	7	5
OEHL	4	1	1
AL	12	10	6
RADC	15	11	8
SAM	17	16	9
WL	8	7	3
WHMC	1	0	0
Total	168	134	96

LIST OF PARTICIPATING UNIVERSITIES

Alabama, University of	-	1	New York, State University of	-	2
Alfred University	-	1	North Carolina State University	-	1
Arkansas-Pine Bluff, Univ. of	-	1	Northern Arizona University	-	1
Auburn University	-	1	Northern Illinois University	-	1
Bethel College	-	1	Northwestern University	-	1
Boston College	-	1	Notre Dame, University of	-	1
Brescia College	-	1	Ohio State University	-	2
California Polytechnic	-	1	Oklahoma, University of	-	3
California State University	-	2	Old Dominion University	-	1
Cincinnati, University of	-	2	Pennsylvania State University	-	1
Denver, University of	-	1	Pittsburgh, University of	-	1
Eastern Kentucky University	-	1	Rhode Island, University of	-	1
Florida Atlantic University	-	1	San Diego State University	-	1
Florida Institute	-	1	San Jose State University	-	1
Florida, University of	-	4	Savannah State College	-	1
Hamilton College	-	1	Scranton, University of	-	1
Harvard University	-	1	Southern Oregon State College	-	2
Illinois Institute of Technology	-	1	Southwest Texas State University	-	1
Illinois-Rockford, University of	-	1	Tennessee State University	-	1
Illinois State University	-	1	Tennessee Technological Univ.	-	1
Indiana-Purdue, University of	-	1	Texas A&M University	-	6
Kansas State University	-	2	Texas Southern University	-	1
Lawrence Technological University	-	1	Texas-San Antonio, University of	-	3
Long Island University	-	1	Transylvania University	-	1
Lowell, University of	-	1	Trinity University	-	1
Massachusetts, University of	-	2	US Naval Academy	-	1
Michigan, University of	-	1	Utah State University	-	1
Minnesota-Duluth, University of	-	2	Utica College	-	1
Mississippi State University	-	2	Vanderbilt University	-	1
Missouri-Rolla, University of	-	1	Washington State University	-	1
Murray State University	-	1	West Virginia University	-	2
Nebraska-Lincoln, University of	-	2	Wisconsin-Platteville, Univ. of	-	1
New Hampshire, University of	-	1	Worchester Polytechnic Institute	-	1
New York Institute of Technology	-	1	Wright State University	-	6

Total   -   94

PARTICIPANTS LABORATORY ASSIGNMENT

AERO PROPULSION AND POWER DIRECTORATE
(Wright-Patterson Air Force Base)

Dr. Jerry Clark
Wright State University
Specialty: Physics

Dr. Frank Gerner
University of Cincinnati
Specialty: Mechanical Engineering

Dr. Thomas Lalk
Texas A&M University
Specialty: Mechanical Engineering

Dr. Baruch Lieber
State University of New York
Specialty: Aerospace Engineering

Dr. William Schulz
Eastern Kentucky University
Specialty: Analytical Chemistry
760-7MG-079 and 210-10MG-095

Dr. Richard Tankin
Northwestern University
Specialty: Mechanical Engineering

ARMAMENT DIRECTORATE
(Eglin Air Force Base)

Dr. Peter Armendarez
Brescia College
Specialty: Physical Chemistry

Dr. Joseph Brown
Mississippi State University
Specialty: Mechanical Engineering

Dr. Roger Bunting
Illinois State University
Specialty: Inorganic Chemistry

Dr. Satish Chandra
Kansas State University
Specialty: Electrical Engineering

Dr. David Cicci
Auburn University
Specialty: Aerospace Engineering

Mr. William Newbold (GSRP)
University of Florida
Specialty: Aerospace Engineering

Dr. Boghos Sivazlian
University of Florida
Specialty: Operations Research

Dr. Steven Trogdon
University of Minnesota-Duluth
Specialty: Mechanics

Mr. Asad Yousuf
Savannah State College
Specialty: Electrical Engineering

ARMSTRONG LABORATORY
(Brooks Air Force Base)

Dr. Robert Blystone
Trinity University
Specialty: Zoology

Dr. Gwendolyn Howze
Texas Southern University
Specialty: Molecular Biology

Dr. Carolyn Caudle-Alexander
Tennessee State University
Specialty: Microbiology

Dr. Harold Longbotham
University of Texas-San Antonio
Specialty: Electrical Engineering

Dr. James Chambers
University of Texas - San Antonio
Specialty: Biochemistry

Dr. Ralph Peters (1987)
Wichita State University
Specialty: Zoology

Dr. Mark Cornwall
Northern Arizona University
Specialty: Human Performance

Dr. Raymond Quock
Univ. of Illinois at Rockford
Specialty: Pharmacology

Dr. Vito DelVecchio
University of Scranton
Specialty: Biochemical Engineering

Dr. Ram Tripathi
University of Texas-San Antonio
Specialty: Statistics

ARNOLD ENGINEERING DEVELOPMENT CENTER
(Arnold Air Force Base)

Dr. Brian Beecken
Bethel College
Specialty: Physics

Dr. Lang-Wah Lee
University of Wisconsin-Platteville
Specialty: Mechanical Engineering

Dr. Stephen Cobb
Murray State University
Specialty: Physics

Dr. Chun Fu Su
Mississippi State University
Specialty: Physics

Dr. John Francis
University of Oklahoma
Specialty: Mechanical Engineering

Dr. Richard Tipping
University of Alabama
Specialty: Physics

Dr. Orlando Hankins
University of North Carolina State
Specialty: Nuclear Engineering

Dr. D. Wilkes
Vanderbilt University
Specialty: Electrical Engineering

AVIONICS DIRECTORATE
(Wright-Patterson Air Force Base)

Dr. David Choate
Transylvania University
Specialty: Mathematics

Dr. Dar-Biau Liu
California State University
Specialty: Applied Mathematics

Dr. R. H. Cofer
Florida Institute
Specialty: Electrical Engineering

Dr. Robert Shock
Wright State University
Specialty: Mathematics

CREW SYSTEMS DIRECTORATE
(Wright-Patterson Air Force Base)

Dr. Thomas Lockwood
Wright State University
Specialty: Toxicology

Dr. Michael Stanisic
University of Notre Dame
Specialty: Robotics

Dr. Ethel Matin
Long Island University
Specialty: Experimental Psychology

Dr. Chi-Ming Tang
State University of New York
Specialty: Mathematics

Dr. Randy Pollack
Wright State University
Specialty: Anthropology

Dr. Ebo Tei
University of Arkansas-Pine Bluff
Specialty: Psychology

Dr. Donald Robertson (1987)
Indiana University of Pennsylvania
Specialty: Psychology

ENGINEERING AND SERVICES CENTER
(Tyndall Air Force Base)

Dr. William Bannister
University of Lowell
Specialty: Organic Chemistry

Dr. Kim Hayes
University of Michigan
Specialty: Environmental Engineering

Dr. Emerson Besch
University of Florida
Specialty: Animal Physiology

Dr. Deborah Ross
University of Indiana-Purdue
Specialty: Microbiology

Dr. Avery Demond
University of Massachusetts
Specialty: Civil Engineering

Dr. Dennis Truax (1987)
Mississippi State University
Specialty: Civil Engineering

Dr. Kirk Hatfield
University of Florida
Specialty: Civil Engineering

Dr. George Veyera
University of Rhode Island
Specialty: Civil Engineering

ELECTRONIC SYSTEMS DIVISION
(Hanscom Air Force Base)

Dr. Stephen Kolitz (1986)
University of Massachusetts
Specialty: Operations Research

Dr. Sundaram Natarajan
Tennessee Technical University
Specialty: Electrical Engineering

FLIGHT DYNAMICS DIRECTORATE
(Wright-Patterson Air Force Base)

Dr. Kenneth Cornelius
Wright State University
Specialty: Fluid Mechanics

Dr. William Wolfe
Ohio State University
Specialty: Engineering

Dr. Arnold Polak
University of Cincinnati
Specialty: Aerospace Engineering

Dr. Lawrence Zavodney
Ohio State University
Specialty: Mechanical Engineering

Dr. Nisar Shaikh
University of Nebraska-Lincoln
Specialty: Applied Mathematics

FRANK J. SEILER RESEARCH LABORATORY
(United States Air Force Academy)

Dr. Robert Granger
US Naval Academy
Specialty: Mechanical Engineering

Dr. Timothy Troutt
Washington State University
Specialty: Mechanical Engineering

Dr. Clay Sharts
San Diego State University
Specialty: Chemistry

Dr. Hung Vu
California State University
Specialty: Applied Mechanics

GEOPHYSICS DIRECTORATE
(Hanscom Air Force Base)

Dr. Phanindramohan Das
Texas A&M University
Specialty: Geophysical Science

Dr. Thomas Miller
University of Oklahoma
Specialty: Physics

Dr. Alan Kafka
Boston College
Specialty: Geophysics

Dr. Henry Nebel
Alfred University
Specialty: Physics

Dr. Charles Lishawa
Utica College
Specialty: Physical Chemistry

Dr. Craig Rasmussen
Utah State University
Specialty: Physics

HUMAN RESOURCES DIRECTORATE
(Brooks, Williams and Wright-Patterson Air Force Base)

Dr. Kevin Bennett
Wright State University
Specialty: Applied Psychology

Mr. John Williamson (GSRP)
Texas A&M University
Specialty: Psychology

Dr. Deborah Mitta
Texas A&M University
Specialty: Industrial Engineering

Dr. Michael Wolfe
West Virginia University
Specialty: Management Science

Dr. William Smith
University of Pittsburgh
Specialty: Linguistics

Dr. Yehoshua Zeevi
Harvard University
Specialty: Electrical Engineering

Dr. Stanley Stephenson
Southwest Texas State University
Specialty: Psychology

Dr. Robert Zerwekh
Northern Illinois University
Specialty: Philosophy

MATERIALS DIRECTORATE
(Wright-Patterson Air Force Base)

Dr. Donald Chung
San Jose State University
Specialty: Material Science

Dr. Kenneth Currie
Kansas State University
Specialty: Industrial Engineering

Dr. Michael Resch
University of Nebraska-Lincoln
Specialty: Materials Science

Dr. James Sherwood
University of New Hampshire
Specialty: Aerospace Mechanics
210-9MG-088 and 210-10MG-098

Dr. Michael Sydor
University of Minnesota-Duluth
Specialty: Physics

OCCUPATIONAL AND ENVIRONMENTAL HEALTH DIRECTORATE
(Brooks Air Force Base)

Dr. Stewart Maurer
New York Institute of Technology
Specialty: Electrical Engineering

ROCKET PROPULSION DIRECTORATE
(Edwards Air Force Base)

Dr. Lynn Kirms
Southern Oregon State College
Specialty: Organic Chemistry

Dr. Mark Kirms
Southern Oregon State College
Specialty: Organic Chemistry

Dr. Faysal Kolkailah
California Polytechnic
Specialty: Mechanical Engineering

Dr. Vittal Rao
University of Missouri-Rolla
Specialty: Control Systems

Dr. Larry Swanson
University of Denver
Specialty: Mechanical Engineering

Dr. Roger Thompson
Pennsylvania State University
Specialty: Engineering Mechanics

ROME LABORATORIES
(Griffiss Air Force Base)

Dr. Charles Alajajian
West Virginia University
Specialty: Electrical Engineering

Dr. Khaja Subhani
Lawrence Tech. University
Specialty: Electrical Engineering

Dr. Ian Grosse
University of Massachusetts
Specialty: Mechanical Engineering

Dr. David Sumberg (1987)
Rochester Institute of Tech.
Specialty: Physics

Dr. Henry Helmken
Florida Atlantic University
Specialty: Physics

Dr. Donald Ucci
Illinois Institute of Technology
Specialty: Electrical Engineering

Dr. Michael Klein
Worcester Poly Institute
Specialty: Physics

Dr. Kenneth Walter (1988)
Prairie View A&M University
Specialty: Chemical Engineering

Dr. William Kuriger
University of Oklahoma
Specialty: Electrical Engineering

Dr. James Wolper
Hamilton College
Specialty: Mathematics

WEAPONS DIRECTORATE
(Kirtland Air Force Base)

Dr. Harry Hogan
Texas A&M University
Specialty: Mechanical Engineering

Dr. Duc Nguyen
Old Dominion University
Specialty: Civil Engineering

Dr. Arkady Kheyfets (1988)
North Carolina State University
Specialty: Mathematical Physics

Dr. Duane Sanders
Texas A&M University
Specialty: Civil Engineering

RESEARCH REPORTS

MINI-GRANT RESEARCH REPORTS

Technical Report Number	Title and Mini-Grant Number	Professor
Volume I		

Rome Laboratories

1	Optimal Design of Finite Wordlength FIR Digital Filters for an Analog Transversal Filter with Tap Weight Circuitry Defects Using Adaptive Modeling 210-10MG-123	Dr. Charles Alajajian
2	Automatic Adaptive Remeshing for Finite Element Reliability Assessment of Electronic Devices 210-10MG-129	Dr. Ian Grosse
3	Ionospherically-Induced Phase Distortion Across Wide-Aperture HF Phased Arrays 210-10MG-047	Dr. Henry Helmken
4	A Study of Interacting Tunneling Units with Possible Application to High Temperature Superconductors 210-10MG-057	Dr. Michael Klein
5	Reduced Bandwidth Binary Phase-Only Filters 210-10MG-052	Dr. William Kuriger
6	Computer Modeling of GaAs/AlGaAs MQW Devices for Optical Properties 210-10MG-107	Dr. Khaja Subhani
7	Fiber Optic Distribution System for Phased Array Antennas 760-7MG-113	Dr. David Sumberg (1987)
8	Continuation Study of a Communications Receiver for Spread Spectrum Signals 210-10MG-067	Dr. Donald Ucci
9	Development of a System to Deposit Thin Films of Titanium Carbide Using Atomic Layer Epitaxy 219-9MG-113	Dr. Kenneth Walter (1988)

Engineering and Services Center

Frank J. Seiler Research Laboratory

Volume II

Phillips Laboratory

Geophysics Directorate

Volume III

Wright Laboratory

1989 USAF-UES RESEARCH INITIATION PROGRAM

Sponsored by the
AIR FORCE OFFICE OF SCIENTIFIC RESEARCH
Conducted by the
Universal Energy Systems, Inc.

FINAL REPORT

CUMULUS PARAMETERIZATION IN NUMERICAL PREDICTION MODELS: A NEW PARCEL-DYNAMICAL APPROACH

Prepared by: Phanindramohan Das

Academic Rank   Professor

Department and University: Department of Meteorology,
                            Texas A&M University

Research Location: Department of Meteorology,
                            Texas A&M University,
                            College Station, Texas
                            77843-3146

P.O. No: S-210-10MG-087

ABSTRACT

The available cumulus parameterization schemes are based on: (i) Kuo's single class of undiluted towers and its modifications; (ii) Arakawa and Schubert's ensemble of cumuli with varying tops under quasi-equilibrium large-scale forcing; (iii) Lindzen's ensemble of cumuli with varying tops as well as bases; and (iv) Betts' systematic adjustment toward a quasi-equilibrium atmosphere. Physically the Arakawa-Schubert scheme is most nearly complete; but it is both difficult to implement and prone to undesirable heating in the lower atmosphere. The physical basis of the Kuo scheme is somewhat obscure, while those of Lindzen, and Betts are ad hoc.

A new concept for a cumulus parameterization scheme is proposed based on the physics and dynamics of an undiluted tower; it will use moisture convergence to determine cloud coverage and compensatory subsidence for cumulus heating. A microphysical scheme will determine precipitation formation and fallout. The scheme will be closely tied to the convective available potential energy (CAPE), appropriately corrected for the negative buoyancy due to condensed water loading.

ACKNOWLEDGMENTS

I. INTRODUCTION

Convective clouds, and the precipitation resulting from them, need to be represented in all weather and climate prediction models, since they dominate the weather of the tropics, as well as that of the temperate region in the spring and summer. However, the horizontal dimension of a convective cell is on the order of a few kilometers, while the weather stations for routine observations are separated by distances on the order of hundreds of kilometers. Obviously it would be a hopeless task to resolve the individual convective clouds by any practical observing system and the numerical weather prediction (NWP) model based on it. The difficulty of the problem is ameliorated by the fact that many convective clouds form simultaneously over a given region, and the change caused by them can be treated as a statistical average leading to a system of parameters developed from the detailed knowledge of the convective, or cumulus, clouds. The procedure envisioned in the development of these parameters is popularly (as well as technically) known as cumulus parameterization.

The methodology of cumulus parameterization in numerical prediction models initially has been developed by meteorologists who have been interested primarily in large-scale processes and guided by their intuition based on rather simple principles. As the computer technology advanced, higher level of sophistication has been progressively introduced in the parameterization schemes, but mainly in terms of intuitive perceptions of how ensembles of cumuli should behave and interact with the large-scale atmosphere: the physics of individual cumuli, when incorporated in these representations, has remained rather rudimentary. The resulting prediction of weather under predominantly convective conditions, while being successful to various degrees, has not attained a high level of reliability in all geographical locations and under all conditions of convective potential.

The forerunners of cumulus parameterization are what are known as convective adjustments (Manabe and Strickler, 1964; Manabe et al.,1965) which aim at adjusting for convective overturning whenever gravitational instability develops in a numerical model simulation. Instead of explicitly invoking processes that occur in a cumulus cloud, these adjustments are effected through numerical algorithms which adjust the vertical distributions of temperature and moisture such that the resulting atmosphere is convectively stable. A run-down of these adjustments and their skill in rainfall prediction have been given by Krishnamurti et al. (1980)

A conscious effort at parametrically incorporating cumulus processes in a dynamical-numerical model, albeit in a very rudimentary manner, is due to Kuo (1965) who treated the deep cumulus cloud as an undiluted tower. Originally intended for use in the simulation of tropical storms which are known to be driven by systems of convective clouds, the Kuo (1965) scheme was received widely for use in general numerical weather prediction (NWP) models. However, it was soon recognized that the scheme severely underestimated precipitation. In addition, the method of heating and moistening used by Kuo (1965) was not based on any physical mechanism but was seen as an ultimate effect that nature would achieve. In a later study, Kuo (1974) changed his cloud model from an undiluted tower to one with entrainment to be specified by two parameters which were to be determined experimentally. In addition, the heating and moistening effects of (large scale) moisture convergence were apportioned by invoking a parameter with no method specified for its determination. The algorithm for the horizontal mixing of the cloud with the environment was also changed.

While there has been little direct impact of the entrainment formulation of Kuo (1974) on later developments, Krishnamurti et al. (1976) were strongly influenced by the new method of horizontal mixing as well as the partitioning of moisture convergence for moistening and heating. As a matter of fact, Krishnamurti et al.

provided, albeit based on an _ad hoc_ assumption, a method for the determination of the partitioning parameter. It also is noteworthy that they changed the method of computing moisture convergence which was treated as concomitant with mass convergence. Krishnamurti _et al._ (1980) further elucidated their scheme. It appears that the formulations of Kuo parameterization currently in use are, to a great extent, combinations of Kuo (1965, 1974) and Krishnamurti, et al. (1980). Later variations on the Kuo scheme, while laden with complex sophistication, have mostly been ignored by operational prediction models.

As already noted, in the Kuo and Kuo-type parameterization schemes the process by which the cloud affects the large-scale properties of the atmosphere is left unspecified. The cloud model itself is internally contradictory: the undiluted tower is talked of as if it were an imporous pipe through which heat and moisture flows from the cloud base to the top, and also as one which mixes the heat and moisture horizontally at all levels. Even with the modified scheme, this contradiction does not get alleviated to a great extent, since the entrainment itself is used more as a device to change the temperature distribution in the cloud than a complete mechanism for horizontal exchange between the cloud and the large-scale atmosphere.

Because of the inconsistencies in the Kuo-type schemes as well as its failure to produce satisfactory forecasts, a need was felt to replace the idea of a single species of clouds as undiluted towers by one in which the cumulus population was assumed to consist of ensembles differentiated by varying bases and tops. Originally envisioned by Ooyama (1971), the cumulus population was endowed to have the same base, and ensembles were characterized by varying tops such that an ensemble with a given top was assumed to detrain into the layer in which the top exists, the clouds in different ensembles supplying heat and moisture to different layers. This idea was further developed by Arakawa and Schubert (1974) who

specified the distribution of the tops of cumuli through an entrainment parameter.

The complex system of equations that result from the formulation of Arakawa and Schubert is closed by making two assumptions: (i) The vertical mass flux in the large scale environment is equal to the total mass flux through the cloud population (the so-called quasi-equilibrium assumption); and (ii) heating of the atmosphere, level for level, is determined by cumulus-induced subsidence; while moistening occurs primarily through detrainment. Precipitation formation in the clouds is treated in terms of cloud models. Including further assumptions on cloud processes, Arakawa and Schubert develop a model which can be considered physically complete, while there may be questions on individual assumptions. However, as opined by Lindzen (1980), the system of equations based on the Arakawa-Schubert model is hard to solve. Interestingly, this complex scheme does not produce results much superior to that produced by the conceptually simpler Kuo schemes.

Apparently frustrated by the heuristic and ad hoc nature of the Kuo-type schemes and the mathematical difficulties of the Arakawa-Schubert formulation, as well as the failure of both to produce consistently satisfactory forecasts, efforts have been directed toward developing models based on the observed behavior of the atmosphere. One such approach, offered by Lindzen (1980), is to divide the atmosphere into two parts, one below and the other above the level of minimum equivalent potential temperature characteristic of the tropical atmosphere. It is then assumed that the mass converging in a layer of the lower part of the atmosphere will rise into the upper part to a level at which the equivalent potential temperature of the layer of origin is matched. The readjustment caused by the consequent overturning should then adjust the atmosphere to a stratification characteristic of the observed equilibrium atmosphere. Interestingly, schemes based on this very simple approach appears to give results comparable to the Kuo schemes as modified by Krishnamurti et al., and the Arakawa-

Schubert methods (Geleyn et al, 1982). However, as admitted by Lindzen, the method is ad hoc ; as a matter of fact, it takes cumulus parameterization backwards in the sense that it is applied more as an adjustment algorithm than through a parametric treatment of the known physics and dynamics of convective clouds.

Another method, developed by Betts (1986) and tested by Betts and Miller (1986), is simultaneously to relax the instantaneous temperature and moisture fields "towards observed quasi-equilibrium thermodynamic structures," with relaxation times dependent on the model characteristics. This approach as that of Lindzen "sidesteps all the details of how the subgrid-scale cloud and mesoscale processes maintain the quasi-equilibrium structure we observe."

The purely empirical approaches of Lindzen as well as Betts, apparently practical for the tropics where quasi-equilibrium conditions are often available, are likely to fail in the extratropical regions where the atmosphere is mostly in a state of transition. In other words, the search for a cumulus parameterization scheme must continue.

Centers of numerical weather prediction, globally, are currently in the process of assessing the potential of the available cumulus parameterization schemes. One striking finding is that the Kuo (1965) parameterization, as modified by Krishnamurti, et al. (1976), simplistic as it is, works quite well compared to the other schemes. An intriguing fact that emerges from a review of the previous studies is that schemes based on quite distinctly disparate assumptions, such as (i) single species of deep cumuli (Kuo, 1965), (ii) ensembles of cumuli with a common base but varying heights of tops (Arakawa and Schubert, 1974) and (iii) ensembles of cumuli with varying heights of bases and tops (Lindzen, 1980), appear to produce essentially similar results overall, although there are differences in detail. A further source of curiosity is that one of these schemes imply entrainment of environmental air into the cloud, while the

others assume undiluted clouds. Apparently it does not matter whether entrainment is included in the scheme, so long as other factors are properly represented, of which the important ones appear to be the vertical fluxes of moisture and mass that are processed by the cumuli, and the precipitation that reaches the ground. Of these the former can be assessed from large-scale processes, while the latter depends on the microphysics of the clouds. An examination of the last factor by Hack, et al., (1984) reveals considerable potential for cloud-physical estimation of precipitation.

II. OBJECTIVES

The principal objectives of the study presented here are (1) to delineate the philosophy of a new cumulus parameterization scheme based on a cloud-physical model which will be intimately tied to the moisture and mass fluxes in the atmosphere and at the same time include an adequate microphysics of clouds so as to make a realistic prediction of precipitation; and (2) to formulate an algorithm based on this philosophy. The need for the above are clarified in the following review of the principles of cumulus parameterization and an analysis of the available observational and numerical model studies.

III. THE PHYSICAL PRINCIPLES OF CUMULUS PARAMETERIZATION

Cumulus parameterization is a component of the more general problem, namely, that of the conversion of latent heat into sensible heat and potential energy, and the redistribution of moisture in the atmosphere. In a convectively unstable atmosphere this conversion proceeds through the activity of the convective clouds the functions of which are not individually resolved by a large-scale observational network. A necessary assumption is therefore that the collective effect of all the convective clouds in a region is reflected in some variable which can be observed on the large scale. For example, Charney (1963, 1969) argued that in the tropics all the vertical

mass flux appears to be due to the cumulus-induced motion. Another such variable is the vertical flux of moisture. As a matter of fact, both mass and moisture flux as well as their concomitant divergences have been used in cumulus parameterization, and experience does not appear to indicate if one works much better than the other. On the other hand, from the point of view of cumulus physics, there is a closer conceptual relation of cumulus activity to large-scale moisture convergence than mass convergence which, as found by Soong and Tao (1980), is related to cumulus motion in a rather complex fashion. Indeed, as observed by these authors, the mass flux through the core of model convective updrafts is about three times the large-scale mass flux. An additional support for the relation between moisture convergence and cumulus activity comes from the study of Reed and Recker (1971) who found close agreement between observed precipitation and the synoptic-scale moisture convergence. Reed and Recker also noted that a substantial amount of moisture convergence occurs above the boundary layer.

In the extratropical regions, the precipitation in some cases has been found to exceed the large-scale convergence (Fritsch et al, 1976; Lin, 1986[*]). However, these analyses refer to mesoscale systems which have definite organization and as such are outside the basic philosophy of cumulus parameterization.

In view of the comments made above, we shall follow Kuo (1965) and start the physical formulation of cumulus parameterization by evaluating the moisture entering a vertical column of the atmosphere, per unit time, by (large-scale) horizontal convergence and surface evaporation as:

$$M_t = - \int_{z_o}^{z_T} \nabla_o(\rho q v) dz + \rho_o C_d (q_s - q_o), \qquad (1)$$

[1*] Quoted by Cotton and Anthes (1989).

where ρ is the air density, q, the specific humidity, v, the horizontal velocity, C_d, the drag coefficient of the surface, and z is the vertical coordinate, with subscript o referring to the lowest model surface, s to the physical earth's surface and T referring to the top of the moist layer . The moisture represented in M_t is disposed of as two final products: (i) area-averaged precipitation reaching ground at the rate R, and (ii) moisture left behind in the atmosphere in the quantity M_t- R. In other words, we can write

$$M_t = R + bM_t, \qquad (2)$$

where we have formally introduced a partitioning parameter, b:

$$b = \frac{M_t - R}{M_t}. \qquad (3)$$

Obviously, b represents the nonprecipitating fraction of the large-scale moisture convergence that is vertically redistributed by cumulus processes. In other words,

$$bM_t = \int_{z_o}^{z_t} \frac{\partial}{\partial t}(\rho q)dz. \qquad (4)$$

Since the net heating of the atmosphere is a result of the conversion of moisture into precipitation, we have

$$LR = \int_{z_o}^{z_t} \frac{\partial}{\partial t}(\rho s)dz, \qquad (5)$$

where L is the latent heat of vaporization, and s ($= c_p T + gz$) is the static energy.

In developing an algorithm out of the above analysis, we must at first find a method for determining b, which then will immediately give R. This will complete the hydrologic enquiry in that we have magnitudes of how much moisture has been converted to heat and how much has been retained in the atmosphere. However, in a

multilevel description of the atmosphere, we must know how this heat and moisture has been vertically distributed. In other words we must determine the vertical profiles of $\partial(pq)/\partial t$ and $\partial(ps)/\partial t$, which will satisfy (4) and (5).

Up to this point the meteorology of the processes by which the latent heat is converted to static energy has remained unidentified. Fundamentally these processes can be divided into two categories: (i) large-scale precipitation under convectively stable conditions and (ii) convective precipitation. Noting further that precipitation is the crucial process we can subdivide convective processes into two categories: (a) shallow convection, which is principally nonprecipitating, causing no net heating of the atmosphere, but redistributing heat and moisture (in this respect, shallow convection acts more like eddy diffusion); and (b) deep convection, which is the principal process by which convective precipitation is caused.

A comment is in order on this dichotomy of convective processes, since it ignores the central idea behind the Arakawa-Schubert scheme. That cumulus activity usually takes the form of a few deep cumuli embedded in a large population of shallow clouds is a common observation, although among the shallow clouds there certainly is a tendency for a distribution in their depths. Likewise, the tops of all deep clouds do not have the same height. However, it must be realized that in a snapshot view the clouds are seen in various states of development. In final analysis, this dichotomy appears to be supported by observations, as can be seen from the analytical study of Ogura and Cho (1973) for the tropics and that of Lewis (1975) for the extratropical region.

In this study we focus on the deep cumuli and discuss parameterization schemes which will represent the physical and dynamical processes inside and around the cloud through the parameters such as its vertical and horizontal scales and lifetime. The knowledge of how these parameters are determined in natural clouds is at best

fragmentary; consequently simple cloud models are used in all cumulus parameterization schemes.

As already remarked, the first effort at deep cumulus parameterization is due to Kuo (1965). The cloud model adopted by him treated the active cloud and its environment as a single system so that there was no need for treating cloud-environment interactions (e.g., entrainment). The temperature (T_c) and specific humidity (q_c) in this model cloud was determined by those in an "undiluted tower", that is by a moist adiabat through the cloud base conditions. The horizontal scale of the cloud was represented by the fractional area, a, occupied by a population of cumuli, which, in turn, was related to cloud lifetime (τ) through

$$M_t = a\, M_c/\tau, \tag{6}$$

where M_c is the moisture processed in a cloud column of unit cross section . One can write

$$M_c = I_\theta + I_q; \tag{7}$$

where

$$I_\theta = \text{moisture used in heating} = \int_{z_b}^{z_t}(c_p/L)(T_c - T)\rho dz, \tag{8}$$

and

$$I_q = \text{moisture used in moistening} = \int_{z_b}^{z_t}(q_c - q)\rho dz. \tag{9}$$

Precipitation rate over the area under the cumuli is determined from

$$R_c = I_\theta /\tau, \tag{10}$$

in which τ is disposed of as the observed duration of its developing stage. The heat realized in the precipitation was vertically apportioned through the relation:

$$\frac{\partial s/\partial t}{\text{moisture processed per unit time}}$$

$$= \frac{\text{heating per unit mass in a layer}}{\text{moisture processed in cloud lifetime}},$$

or

$$\frac{\partial s}{\partial t} = \frac{c_p(T_c - T)}{M_c} M_t = \frac{a}{\tau} c_p(T_c - T). \qquad (11)$$

with a similar expression for moistening, that is, for the local rate of increase of q with time.

The procedure set by Kuo (1965) does not explicitly determine b but it can be easily shown to imply

$$b = \frac{a}{\tau} \frac{l_q}{M_t}. \qquad (12)$$

Obviously, this expression for b has no physical basis. Later Kuo (1974) recognized, in addition, that his 1965 parameterization severely underestimated rainfall. He therefore suggested that an explicit empirical determination of b be made but gave no method. It was Krishnamurti et al. (1976) who provided a method which was elucidated further by Krishnamurti et al. (1980). In principle, it differs little from that implied in (12); in actual application, however, b is expressed as the ratio of the columnar integral of horizontal moisture advection to that of the three-dimensional moisture divergence. As recognized by the authors, the method, while giving realistic prediction in the cases tested, has little physical basis. As a matter of fact there is no direct physical basis for determining b from these approaches which therefore must be abandoned in a physical model.

In the model of Arakawa and Schubert (1974) precipitation is treated directly but in terms of highly simplified cloud physics.

Hack et al (1984) have made an effort to improve their algorithm by adapting a parametric model of Lopez (1973) for cumulus precipitation. In the model to be presented in this study, it is the precipitation (R) which will be the primary hydrologic parameter to be obtained from the cumulus algorithm. Once R is known, b will follow from (3).

The methods of determination of the vertical profiles of $\partial s/\partial t$ and $\partial q/\partial t$ needed for completing the parameterization, as given by Kuo, Anthes (1977), and Krishnamurti and collaborators, also have no direct physical basis. An obviously appropriate method is to apply another aspect of the physics of the Arakawa-Schubert model, apparently first presented by Arakawa (1969), namely, the use of cumulus-induced subsidence to determine the heating and moistening of the environment. This approach is soundly based on the studies of Bjerknes (1938) and Asai and Kasahara (1967).

IV. THE PHILOSOPHY OF THE NEW PARAMETERIZATION SCHEME:

The cumulus parameterization scheme being presented in this study, is purported to include a view of the physics and dynamics of convective clouds as developed by cumulus microdynamicists. The basic philosophy of the scheme is that of the Kuo-type ones in that it adopts moisture convergence rather than the vertical mass flux as the factor determining the horizontal coverage of cumulus activity. The support for this view point, while not unequivocal, is quite strong (Ogura and Cho,1974; Reed and Recker, 1971). The mode of interaction of the cloud with the large-scale atmosphere, however, is different from that the Kuo-type parameterization: it is the cumulus-induced mass flux which, through compensatory subsiding motion, plays a dominant role in the vertical redistribution of heat and moisture.

The scheme, as already emphasized, will consider only deep convection, and while resorting to some minor assumptions of an ad hoc

nature, will be reasonably well tied to a simple cloud model. The physic assumptions underlying the scheme are as follows:

a. Single class of undiluted towers of vertically variable radius: In view of the apparent success of the Kuo-type schemes as well as the comments made by Lindzen (1980), a complex cloud model may not be warranted. Since the products of the cloud are ultimately mixing with the large-scale environment entrainment of environmental air need not be explicitly considered (Kuo,1965). However, entrainment has strong effect on buoyancy and vertical velocity, but it is crucially dependent on the horizontal dimension of the cloud, which, in turn, affects the perturbation pressure and, through it, the vertical velocity in the cloud. In short, the horizontal dimension of the cloud has too complex an effect to be considered in a preliminary effort. The importance of entrainment is further minimized by the study of Kuo and Anthes (1984), which shows that the vertical profiles of heating computed with no entrainment "are very similar to those computed from a large cloud radius with entrainment and are very close to the observed profile."

b. Simultaneously occurring clouds in different parts of their life cycle: In the Arakawa-Schubert and Lindzen schemes, the need for clouds of various depths has been recognized, apparently for the purpose of modifying all the layers of the atmosphere by the cumulus processes. Unfortunately, there is no physically established distribution function for cloud depths. On the other hand, even for a single species of deep clouds, there are cloud cells in different stages of their life cycles. Thus the following cloud events will be considered as simultaneously occurring but for different clouds: (a) developing clouds, signifying upward mass flux and moisture convergence; (b) dissipating clouds, signifying falling precipitation, evaporation of precipitation in the lower part and mixing of the cloud with the environment in the upper part; and (c) mature clouds, having the characteristics of developing clouds in the uppermost layers and dissipating clouds in the lowermost layers. The cloud model consistent with the above physical picture will consist of a

one-dimensional, steady-state, nonentraining parcel model for the developing phase and a precipitating model for the mature and the dissipating phase.

c. **The developing cloud.** The depth of the developing cloud will be determined by the convective available potential energy (CAPE), the computation of which starts with the computation of the convective condensation level (CCL) or the level of free convection (LFC). It is not clear, at present, which is the preferable parameter, since there is no clear understanding of the mode of initiation of convection. A study by Petterssen *et al.* (1945) shows a high correlation of the cloud base with the CCL. In the present study the mixing ratio at the CCL is determined by the average mixing ratio in the lowest 5-kPa layer of the atmosphere. The computation of CAPE will include a correction for the negative buoyancy due to the condensed phase (Das, 1964; Seitter and Kuo, 1983), or, in other words, a "cloud-virtual temperature" correction of Saunders (1957), the CAPE will be expressed as

$$CAPE = g \int_{z_b}^{z_m} \frac{1}{T_v}(T_{vc} - T_v)dz, \qquad (13)$$

where z_b is the height of the CCL, z_m refers to the top of the "positive area" and the other symbols are: g, the acceleration due to gravity; T_{vc}, the cloud virtual temperature and T_v is the virtual temperature in the environment. T_v and T_{vc} can be expressed as

$$T_v = T (1 + 0.61q), \text{ and } T_{vc} = T_c (1 + 0.61q_c - q_\ell), \qquad (14)$$

where T_c is the temperature on the moist adiabat passing through the characteristic point at the CCL, and q_c, the saturation mixing ratio at T_c and the pressure at the relevant level. The liquid-water mixing ratio, q_ℓ, in the adiabatic cloud is obtained as the difference

between the q_b at the CCL and q_c at the level under consideration, that is,

$$q_t = q_b - q_c.$$ (15)

d. __Fractional area covered by the clouds.__ The fractional area covered by the cloud will vary with height and will be determined so as to have the available moisture convergence to pass vertically through the developing cloud at the level of maximum vertical velocity in the cloud. At other levels, the cloud cover will be given by the assumption of constant mass flux through the cloud. This assumption is based on the possibility that most of the air ascending through the cloud comes from the lowest layers of the atmosphere.

The upward mass flux through the cloud will be balanced, layer for layer, by compensatory downward flux in the environment so that there is no cumulus-induced large-scale pressure change (Bjerknes, 1938; Asai and Kasahara, 1967). The implication of this assumption is that large-scale mass flux is not __directly__ affected by the cumulus-induced vertical motion. Indirectly, however, the large-scale vertical mass flux will be affected through the net increase of the static energy and its effect on the geopotential of the atmosphere.

e. __Precipitation formation and fallout__: The precipitation formation will be associated with the mature phase, and will be determined by a steady-state solution of a cloud-to-precipitation conversion equation, with a conversion parameter analogous to that of Hack, __et al.__ (1984). The use of such a conversion parameter is an oversimplification; however, in a preliminary effort, it will avoid the complexity of microphysical processes envisioned by cumulus physicists.

Precipitation fallout will be associated with the dissipating stage of the cumulus, and will be assumed to fall in stationary air. No

downdraft formation is included in the preliminary formulation and application of the scheme. The rationale for this approach is that the vertical mass balance function of the downdraft is obviated by the layer-by-layer application of the compensatory subsidence mentioned above. On the other hand, the transport of precipitation by the downdraft is a local process and is not likely to have importance in determining the (average) precipitation on large space-time scales. The evaporation of the falling precipitation and the consequent cooling of the environmental air will be included however.

f. Large-scale heating. Large-scale heating will be effected primarily by compensatory subsidence in the environment of the cloud. This heating will be modulated by the mixing of the dissipating cloud with the environment inasmuch as there is cooling due to evaporation of residual cloud water and heating (or cooling) due sensible heat excess (or deficit) of the cloud. In addition, in the lower part of the cloud, cooling due to the evaporation of precipitation will be included.

The physics of the dissipation of the cloud and the way it mixes with the environment is not clear. The only thing that can be definitely said is that the mixing occurs through turbulent diffusion which is enhanced by the motion induced by the cumulus activity itself. It is the opinion of the author is that any mixing algorithm will be ad hoc. In a preliminary effort, therefore, it is suggested that a simple algorithm be adopted and the most appealing one, in that respect, appears to be that due to Kuo (1965). This apparently runs counter to Kuo and Anthes (1984) who find it necessary to use a complex eddy sensible heat flux algorithm. It should be remembered, however, that the philosophy of parameterization suggested here is quite different from that of these authors. More detail of the algorithm will be presented in the next section.

g. **Large-scale moistening**. Large-scale moistening due to cumuli will be through the mixing algorithm discussed above for the heating case.

V. PHYSICS AND MATHEMATICS OF A SAMPLE ALGORITHM:

The algorithm presented below assumes that large-scale values of temperature, T, mixing ratio, q, pressure, p, and wind velocity v at each grid point of the model domain have been obtained either from observations or from an earlier model run.   The step-by-step algorithm below is presented in height (z) coordinates but can be readily translated into pressure or sigma coordinates.

a. **Computing the developing-cloud parameters**. Even before determining if a deep convective cloud will develop in a given large-scale situation, one needs to determine the conditions such a cloud would have in case it did develop. It may be recalled that our model of a developing-cloud is that of an undiluted, nonprecipitating tower. The base of this cloud will have the height, z_b, and pressure, p_b, which are the same as those of the convective condensation level (CCL).   The temperature at this height is T_b; the mixing ratio in the cloud is q_b which, by definition, is the same as the average mixing ratio in the lowest 5-kPa layer of the atmosphere, which also is the saturation mixing ratio at the temperature T_b and pressure p_b. The temperatures, T_c, and mixing ratios, q_c, inside the cloud are those on the moist–adiabat passing through the characteristic point (T_b, p_b). There are numerous algorithms available for obtaining the above parameters and they will not be detailed here.

The first step in determining if a deep cloud will develop is to determine the CAPE as given by (13) along with (14) and (15) through a simple numerical integration procedure.   However, in determining z_m which should have the conditions

$$\text{CAPE} > 0; \ z_b < z = z_m: \ T_{vc} = T_v; \ \underline{\text{and}} \ z > z_m: \ T_{vc} < T_v, \qquad (16)$$

care is needed since the moist adiabat through the cloud base characteristic point may intersect the temperature sounding at more than one point.

Once a (positive) CAPE and z_m is obtained according to (16), the follwing parameters are readily obtained:

(i) Height of the top of the model developing cloud, z_t: z_t is numerically retrieved from the condition

$$CAPE + g \int_{z_m}^{z_t} \frac{1}{T_v}(T_{vc} - T_v)dz = 0. \qquad (17)$$

(ii) Cloud updraft and its distribution with height. For steady-state, one-dimensional motion of a model parcel, the vertical momentum equation can be written

$$\frac{dw_c}{dt} = w_c \frac{dw_c}{dz} = \frac{g}{T_v}(T_{vc} - T_v) - \frac{1}{\rho}\frac{dp'}{dz},$$

where w_c is the vertical velocity of the cloud parcel and p' is the deviation of the parcel pressure from that of the environment. This pressure deviation, while being important, cannot be directly accounted for in a one-dimensional framework. However, certain idealization permits an approximate correction for the pressure deviation by reducing the buoyancy term (Das, unpublished manuscript). For a cloud with comparable vertical and horizontal dimensions the reduction factor is heuristically seen to be 2/3, so that the last equation can be written

$$\frac{dw_c^2}{dz} = \frac{4}{3}\frac{g}{T_v}(T_{vc} - T_v).$$

This can be formally integrated to give

$$w_c{}^2(z) - w_b{}^2 = \frac{4g}{3} \int_{z_m}^{z_t} \frac{1}{T_v}(T_{vc} - T_v)dz = I_1(z), \quad \text{say},$$

where w_b is the parcel velocity at the cloud base. Constraints on the choice of w_b will be discussed later. In the meantime, we have

$$w_c(z) = [w_b{}^2 + I_1(z)]^{1/2}. \tag{18}$$

(iii) **Mass and moisture flux through the cloud.** At $z = z_m$, the top of the positive area determining the CAPE, w_c will have the highest magnitude, w_m, which will be used to determine the vertical flux densities of mass (M_c) and moisture (water vapor **plus** liquid water) (M_{qc}) through the cloud:

$$\text{(a)} \ M_c = \rho_m w_m; \quad \text{(b)} \ M_{qc} = (q_c + q_\ell)\rho_m w_m \tag{19}$$

where the subscript m refers to the values at $z = z_m$. If, at this height, A_m be the horizontal area of the cloud, we shall have by the assumptions made earlier:

(a) Cloud mass flux $= A_m M_c = A_m \rho_m w_m = $ constant $= A_c \rho w_c$; (20a)

(b) Cloud moisture flux $= A_m M_{qc} = A_m(q_c + q_\ell)\rho_m w_m$

$$= \text{constant} = A(q_c + q_\ell)\rho w_c, \tag{20b}$$

the last members of the above equalities referring to arbitrary heights within the clouds. It should be noted that A_c, the area of the horizontal section of the cloud, varies with height.

b. **Large-scale moisture convergence and fractional area covered by clouds.** As assumed, the vertical flux of moisture induced by the ...ge-scale horizontal moisture convergence will pass through the cloud. We shall write M_q for the areal density of the moisture flux and express it as

$$M_q = - \int_{z_0}^{z_m} \nabla_0(\rho q v) dz + \rho_0 C_d (q_s - q_0), \qquad (21)$$

which is somewhat different from the M_t written in (1). The vertical moisture flux over numerical mesh area A_n obviously will be $A_n M_q$ which is assumed to be equal to $A_m M_{qc}$, so that the fraction of the mesh area convered by the cloud at $z = z_m$ will be

$$a_m = A_m/A_n = M_q/M_{qc}, \qquad (22)$$

and the fractional cloud cover at an arbitray z is

$$a_c = a_m \rho_m w_m/\rho w_c. \qquad (23)$$

With the above evaluation of CAPE, z_t, M_q, w_m, a_m, w_c, and a_c, we are ready to obtain several important pieces of information as follows:

(i) <u>Decision if deep convective clouds will form.</u> There are three obvious parametric thresholds for the formation of the deep convective clouds which can be expressed through the conditions:

$$\text{(a) } a_m \geq a_1; \text{ (b) } w_m \geq w_1; \text{ (c) } z_t \geq z_1,$$

where a_1, w_1, and z_1 are the threshold values of the fractional cloud cover, maximum cloud updraft, and height of the cloud top, respectively. As examples, we can arbitrarily set $a_1 = 10^{-3}$, $w_1 = 5$ m s^{-1}, and $z_1 = 5$ km. Unfortunately, any such set of values has not been systematically established from observations, and consequently, must be obtained by experimentation.

(ii) <u>Cumulus-induced subsidence.</u> Assuming, after Bjerknes (1938), that the mass flux through the cloud is balanced, layer for layer, by subsiding motion in the environment, we can write a one-dimensional equation of continuity:

$$a_c w_c + (1 - a_c) w_e = 0, \qquad (24)$$

where w_e is the velocity of subsidence. It is readily seen that (24) leads to

$$w_e = -\frac{a_c w_c}{1 - a_c}. \tag{25}$$

This w_e along with w_c will be used later to determine cumulus induced heating and moistening.

(iii) <u>Cloud-base vertical velocity.</u> So far we have left w_b unspecified. The reason is easily seen by reference to (23) since if $w_c(z_b) = w_b = 0$, $a_c(z_b)$ becomes infinite, while it should be less than unity, about 10% being a realistic value. In other words, if w_m is about 30 m s^{-1}, with $a_m \approx 0.01$, we shall need to have $w_b \approx 1.5$ m s^{-1} (assuming $\rho_m/\rho_b \approx 0.5$, where ρ_b is the air density at the cloud base). It appears that the popular choice of $w_b \approx 1.0$ m s^{-1} should give reasonable result.

c. <u>Partitioning of condensation products into cloud and rain.</u> The physical algorithm presented below does not include ice phase but can easily be extended to do so. The process envisioned appears relevant to the early mature stage of the cloud when the updrafts are not materially diminished from the developing phase but precipitation formation has progressed sufficiently.

The problem of precipitation formation can be formulated through the following statement: the total nonprecipitating moisture flux at the level z + dz is that at the level z <u>minus</u> that part of cloud water coverted to precipitation within the volume $a_c$dz. Assuming no storage of cloud water in the cloud a conservation equation for water substances can be written

$$[a_c \rho w_c (q_c + q_\ell)]_{z+dz} = [a_c \rho w_c (q_c + q_\ell)]_z - a_c \rho \alpha (q_\ell - \ell_{crit})dz, \tag{26}$$

where $\mathfrak{L}_{crit}$ is a critical cloud-water content which needs to be exceeded for the conversion of cloud water to precipitation to be possible (Kessler, 1969), α being the conversion factor. Since $a_c \rho w_c$ is constant with height, (26) can be rewritten as

$$\frac{\partial q_{\mathfrak{t}}}{\partial z} + \frac{\alpha}{w_c}(q_{\mathfrak{t}} - \mathfrak{L}_{crit}) = -\frac{\partial q_c}{\partial z}. \qquad (27)$$

In the above, following Hack et al. (1984), we are using α to combine the effects of autoconversion as well as accretion. The latter effect is much greater in deeper clouds than the shallower ones. This can be accounted for by making α to increase with cloud depth in a manner done by Hack et al. A value of 0.5 g m^{-3} for $\mathfrak{L}_{crit}$ should give fairly satisfactory results.

(i) <u>Steady-state cloud water profile</u>. Equation (27) can be solved either analytically or numerically to obtain a vertical distribution of $q_{\mathfrak{t}}$ in the cloud. This is the residual cloud water which, in the dissipating stage, will mix with the environment.

(ii) <u>Evaluation of precipitation fallout including evaporation</u>. The physical process envisioned in the fallout of precipitation is conceptually complex although quite straightforward in application. As already mentioned, precipitation formation proceeds in the developing and mature clouds, while the fallout occurs from the late mature and dissipating clouds. It should be realized that precipitation formation should more or less end when the cloud enters the dissipating stage. Thus precipitation formation in the developing and mature clouds is considered simultaneous with the precipitation fallout from the dissipating clouds which ordinarily are characterized by a cloud-water envelop at the upper levels and rain falling below through cloud-free air. The fallout picture envisioned is as follows: all the precipitation formed above a given level will fall through that level steadily except for evaporation in the ambient air. Evaporation itself is treated in two ways: at the levels $z \geq z_m$ precipitation is assumed to fall through saturated,

cloudy air so that no evaporation occurs, while at the levels $z < z_m$ rain evaporates as if it is falling through ambient air. This description leads to the following formal expressions:

$$M_p(z_m) = \int_{z_m}^{z_t} a_c \rho \alpha (q_\ell - \ell_{crit}) dz,$$

where $M_p(z_m)$ denotes the vertical large-scale flux of precipitation falling through the level $z = z_m$; and the large-scale precipitation reaching surface ($z = z_0$):

$$M_p(z_0) = M_p(z_m) \left[1 - \int_{z_0}^{z_m} \beta(q_s - q) dz \right]$$

$$+ \int_{z_0}^{z_m} a_c \rho \alpha (q_\ell - \ell_{crit}) \left[1 - \int_{z_0}^{z} \beta(q_s - q) dz' \right] dz,$$

where q_s is the saturation mixing ratio at the ambient temperature and pressure, q is the ambient mixing ratio, and β is an evaporation parameter per unit height of fall of precipitation. The above expressions are much more easily visualized in the finite-difference form. Writing k for the index of vertical grid points and letting k increase with height, we have the iterative expression

$$M_p(k) = M_p(k + 1) - (\Delta EVAP)_k + (\Delta PROD)_k,$$

where

$$(\Delta EVAP)_k = \beta(q_s - q)\Delta z M_p(k + 1), \text{ and}$$

$$(\Delta PROD)_k = [a_c \rho \alpha (q_\ell - \ell_{crit})]_k \Delta z.$$

For obtaining surface precipitation the iteration is continued down to the corresponding value of k, say, k_0. Then surface precipitation is obtained as

$$R = M_p(k_0)/\delta_\ell,$$

where δ_ℓ is the density of water (1000 kg m^{-3}). It should be noted that the precipitation obtained through the algorithm is on the grid scale and not on the cloud scale.

d. **Cumulus heating and moistening.** The thermodynamic and water conservation equations for the large-scale model can be written respectively as

$$\frac{\partial \theta}{\partial t} + v_0 \nabla \theta + w \frac{\partial \theta}{\partial z} = (\frac{\partial \theta}{\partial t})_c + Q_h,$$

and

$$\frac{\partial q}{\partial t} + v_0 \nabla q + w \frac{\partial q}{\partial z} = (\frac{\partial q}{\partial t})_c + Q_m,$$

where Q_h and Q_m denote additional effects due to eddy mixing and diabatic heating. The purpose of cumulus parameterization is to evaluate $(\frac{\partial \theta}{\partial t})_c$ and $(\frac{\partial q}{\partial t})_c$, which, in the following, will be referred to as "heating" and "moistening."

(i) **Heating.** Heating due to cumulus activity proceeds differently in the layers above and below $z = z_m$. Thus

$$z \geq z_m : (\frac{\partial \theta}{\partial t})_c = -(1- a_c)w_e \frac{\partial \theta}{\partial z} - \frac{L}{c_p} a_c \frac{q_\ell}{\tau} ; \text{ and}$$

$$z \leq z_m : (\frac{\partial \theta}{\partial t})_c = -(1- a_c)w_e \frac{\partial \theta}{\partial z} - \frac{L}{c_p}[a_c \frac{q_\ell}{\tau} + (\Delta EVAP)].$$

In the above we have introduced a time parameter τ which can be interpreted as the duration of the life of cumulus cell. In practice, it needs to be determined by experimentation.

(ii) <u>Moistening</u>. The process is very similar to that for heating. Thus

$$z \geq z_m : (\frac{\partial q}{\partial t})_c = -(1-a_c)w_e\frac{\partial q}{\partial z} + a_c\frac{q_c + q_\ell - q}{\tau} \; ; \; \text{and}$$

$$z \leq z_m : (\frac{\partial q}{\partial t})_c = -(1-a_c)w_e\frac{\partial q}{\partial z} + a_c\frac{q_c + q_\ell - q}{\tau} + (\Delta EVAP).$$

VI. RECOMMENDATIONS

The physical-mathematical structure of a new cumulus parameterization algorithm has been completely developed. The next step recommended is to implement this algorithm, experiment with the disposable parameters and test the realism of the results semiprognostically with data sets collected in large-scale experiments, such as the GATE (GARP Atlantic Tropical Experiment).

Once the semiprognostic tests produce satisfactory results, the computer codes of the Geophysics Laboratory's global prediction model should be modified, and the cumulus parameterization scheme be tested in full-scale prediction mode.

REFERENCES

Anthes, R. A., 1977: A cumulus parameterization scheme utilizing a one-dimensional cloud model. Mon. Wea. Rev., 105, 270-286.

Arakawa, A, and W. H. Schubert, 1974: Interaction of a cumulus cloud ensemble withe the large-scale environment. Part I. J. Atmos. Sci. 31, 674-701.

Asai, T., and A. Kasahara, 1967: A theoretical study of the compensating downward motions associated with cumulus clouds. J. Atmos. Sci. 24, 487-496.

Betts, A. K., 1986: A new convective adjustment scheme. Part I: Observational and theoretical basis. Quart. J. Roy. Meteor. Soc., 112, 677-691.

Betts, A. K., and M. J. Miller, 1986: A new convective adjustment scheme. Part II: Single column tests using GATE wave, BOMEX, ATEX and arctic air-mass data sets. Quart. J. Roy. Meteor. Soc., 112, 693-709.

Bjerknes, J., 1938: Saturated-adiabatic ascent of air through dry-adiabatically descending environment. Quart. J. Roy. Meteor. Soc., 64, 325-330.

Cotton, W. R., and R. A. Anthes, 1989: Storms and Cloud Dynamics. Academic Press, 880 pp.

Das, P., 1964: Role of condensed water in the life cycle of a convective cloud cloud. J. Atmos. Sci. 21, 404-418.

Fritsch, J. M., C. F. Chappell, and L. R. Hoxit, 1976: The use of large-scale for convective parameterization. Mon. Weather Rev., 104, 1408-1418

Geleyn, J.-F., C. Girard, and J.-F. Louis, 1982: A simple parameterization of moist convection for large-scale atmospheric models. Beitr. Phys. Atmosph., 55, 325-334.

Hack, J. J., W. H. Schubert, and P. L. Silva Dias,1984: A spectral parameterization for use in numerical models of the tropical atmosphere. Mon. Wea. Rev, 112, 704-716.

Kessler, E., 1969: On the distribution and continuity of water substance in atmospheric circulations. Meteor. Monogr., 10, No. 32, 84 pp.

Krishnamurti, T. N., M. Kanamitsu, R. Godbole, C.-B. Chang, F. Carr, and J. H. Chow, 1976: Study of a monsoon depression (II), dynamical structure. J. Meteor. Soc. Japan, 54, 208-226.

Krishnamurti, T. N., Y. Ramanathan, H.-L. Pan, R. J. Pasch, and J. Molinari, 1980: Cumulus parameterization and rainfall rates I. Mon. Wea. Rev, 108, 465-472.

Kuo, H.-L., 1965: On formation and maintenance of tropical cyclones through latent heat release by cumulus convection. J. Atmos. Sci., 22, 40-63.

Kuo, H.-L., 1974: Further studies of the parameterization of the influence of cumulus convection on large-scale flow. J. Atmos. Sci., 31, 1232-1240.

Kuo, Y.-H., and R. A. Anthes, 1984: Semiprognostic tests of of Kuo-type cumulus parameterization schemes in extratropical convective system. Mon. Wea. Rev., 112, 1498-1509.

Leslie, L. M., 1981: Comparative performance of convective parameterization schemes in a short-term prognosis model. Beitr. Phys. Atmosph., 54, 173-185.

Lewis, J. M., 1975: Tests of Ogura-Cho model on a prefrontal squall-line case. Mon. Wea. Rev., 103,764-778.

Lin, M.-S., 1986: The evolution and structure of meso-α-scale convective complexes by composite analysis. Ph.D. Thesis, Colorado State University.

Manabe, S., and R. F. Strickler, 1964: On the thermal equilibrium of the atmosphere with a convective adjustment. J. Atmos. Sci., 21, 361-385.

Manabe, S., J. Smagorinsky, and R. F. Strickler, 1965: Simulated Climatology of a general circulation model with a hydrologic cycle. Mon. Wea. Rev. 93, 769-798.

Ogura, Y., and H.-R. Cho, 1973: Diagnostic determination of cumulus cloud population from large-scale variables. J. Atmos. Sci., 30, 1276-1286.

Ogura, Y., and H.-R. Cho, 1974: On the interaction between the subcloud and cloud layers in tropical regions. J. Atmos. Sci., 31, 1850-1859.

Ooyama, K., 1971: A theory on parameterization of cumulus convection. J. Meteor. Soc. Japan, 39, 369-379.

Petterssen, S., E. Knighting, R. W. James, and N. Herlofson,1945, Geofys. Publikasioner. Norske Videnskaps-Akad. Oslo. 16, No. 10.

Reed, R. J., and E. E. Recker, 1971: Structure and properties of synoptic-scale wave disturbances in the equatorial western Pacific. J. Atmos. Sci., 28, 1117-1133.

Saunders, P. M., 1957: Thermodynamics of saturated air: a contribution to classical theory. *Quart. J. Roy. Meteor. Soc.*, **83**, 342-350.

Seitter, K. L., and H.-L. Kuo, 1983: The dynamic structure of squall-line type thunderstorms. *J. Atmos. Sci.*, **40**, 2831-2854.

Soong, S.-T., and W.-K. Tao, 1980: Response of deep tropical cumulus clouds to mesoscale processes. *J. Atmos. Sci.*, **37**, 2016-2034.

Yanai, M., S. Esbensen, and J.-H.Chu, 1973: Determination of bulk properties of tropical cloud clusters from large-scale heat and moisture budgets. *J. Atmos. Sci.*, **30**, 611-627.

Rg AS A DEPTH DISCRIMINANT FOR
EARTHQUAKES AND EXPLOSIONS
IN NEW ENGLAND AND EASTERN KAZAKHSTAN

Alan L. Kafka
Associate Professor of Geophysics
Weston Observatory
Dept. of Geology and Geophysics
Boston College
Weston, MA 02193

FINAL REPORT
Research Initiation Program Grant

ABSTRACT

Seismograms of quarry and mining blasts recorded near the Kazakhstan nuclear test site in the Soviet Union were analyzed and compared with seismograms recorded from earthquakes and explosions in New England. The Soviet data were recorded by stations operated in 1987 by the National Resources Defense Council (NRDC), and the New England data were recorded by the New England Seismic Network (NESN, operated by Weston Observatory). The NRDC seismograms were processed to simulate the NESN instrument response. Thus, seismograms recorded in eastern Kazakhstan were compared with New England seismograms by investigating the eastern Kazakhstan blasts as they would have been recorded if NESN type instruments had been operating in that area.

A method for using short-period Rayleigh waves (Rg) as a depth discriminant in New England has been developed by the Principal Investigtor. That method involves estimating Rg/Lg ratios by using a narrow bandpass filter analysis to measure amplitudes of the Rg and Lg phases at appropriate frequencies and arrival times for each phase. The purpose of this research is to evaluate the extent to which the Rg/Lg ratio method, developed for New England, can be applied to seismograms recorded at local and regional distances from the Kazakhstan test site. The long-term goal of this type of research is to develop seismic methods for monitoring nuclear test ban treaties.

1. INTRODUCTION

This report presents the results of a study of seismograms recorded in the vicinity of the Soviet nuclear test site in eastern Kazakhstan. Seismograms of quarry and mining blasts recorded near the Kazakhstan test site (KTS) were analyzed and compared with seismograms recorded from earthquakes and explosions in New England. The New England data were recorded by the New England Seismic Network (NESN, operated by Weston Observatory). The purpose of this research is to evaluate the extent to which results of studies of short-period Rayleigh waves (Rg) as a depth discriminant in New England can be applied to seismograms recorded at local and regional distances from the KTS. The long-term goal of this type of research is to develop seismic methods for monitoring nuclear test ban treaties.

Seismic waves generated by nuclear explosions and earthquakes travel through the Earth's interior and are recorded at significant distances from the seismic event. The analysis of seismograms of earthquakes and explosions provides information about whether a seismic event was an earthquake or an explosion. If an event is identified as an explosion, seismograms can be analyzed to provide estimates of the yield. Seismic monitoring therefore plays an important role in the surveillance of nuclear weapons tests (e.g. Office of Technology Assessment, Report on Seismic Verification, 1988).

One of the basic problems associated with monitoring a comprehensive nuclear test ban treaty is that of discriminating between small explosions and earthquakes based on local and regional seismic data. Chemical explosions are used routinely in the mining and construction industries in both the U.S. and the Soviet Union. Although the specific depths of these industrial explosions are often unknown, they usually occur at very shallow depths (a few tens of feet), and probably are all shallower than a few hundred meters. Most nuclear explosions are detonated at depths of less than about one kilometer, and the deepest underground nuclear explosions are on the order of a few kilometers deep. In contrast, many earthquakes occur much deeper in the Earth's crust. Thus, accurate estimation of depths of seismic sources is one aspect of discriminating earthquakes from explosions.

Until recently, the Soviet Union has been very restrictive about allowing western seismologists to analyze seismograms recorded by instruments located within its borders. Thus, U.S. seismic monitoring of nuclear tests had to be based on seismic recordings obtained from stations outside of the Soviet Union. In 1986, the Soviet Union permitted a private group from the U.S., the National Resources Defense Council (NRDC), to install the first U.S. controlled seismic stations within the Soviet Union, near the KTS (Figure 1). More recent agreements between the U.S. and the Soviet Union have paved the way for a number of permanent seismic stations on Soviet soil that are jointly operated by U.S. and Soviet scientists (e.g. Simarski, 1991). In this research, I analyzed data recorded in 1987 by the NRDC stations.

2. OBJECTIVES

The Principal Investiagtor (PI) for this research project has been studying the use of Rg as a depth discriminant for seismic sources recorded in New England (e.g. Kafka, 1990). Using seismic data recorded by the NESN, we have been using New England as a "laboratory" for Rg wave propagation in various geological regions. The data recorded in eastern Kazakhstan by the NRDC made it possible to compare Rg waves in New England with Rg waves in eastern Kazakhstan. This research project had two objectives:

- To compare Rg wave excitation and propagation in New England with that of Rg waves recorded in eastern Kazakhstan

- To apply the results of studies of Rg waves as a depth discriminant in New England to seismograms recorded in the Soviet Union.

3. CHARACTERISTICS OF Rg SIGNALS RECORDED IN THE VICINITY OF THE KAZAKHSTAN TEST SITE

One of the major difficulties in using seismic methods to monitor nuclear testing is that seismograms are recordings of ground motion at the recording site, usually located at some distance from the source. The properties of the
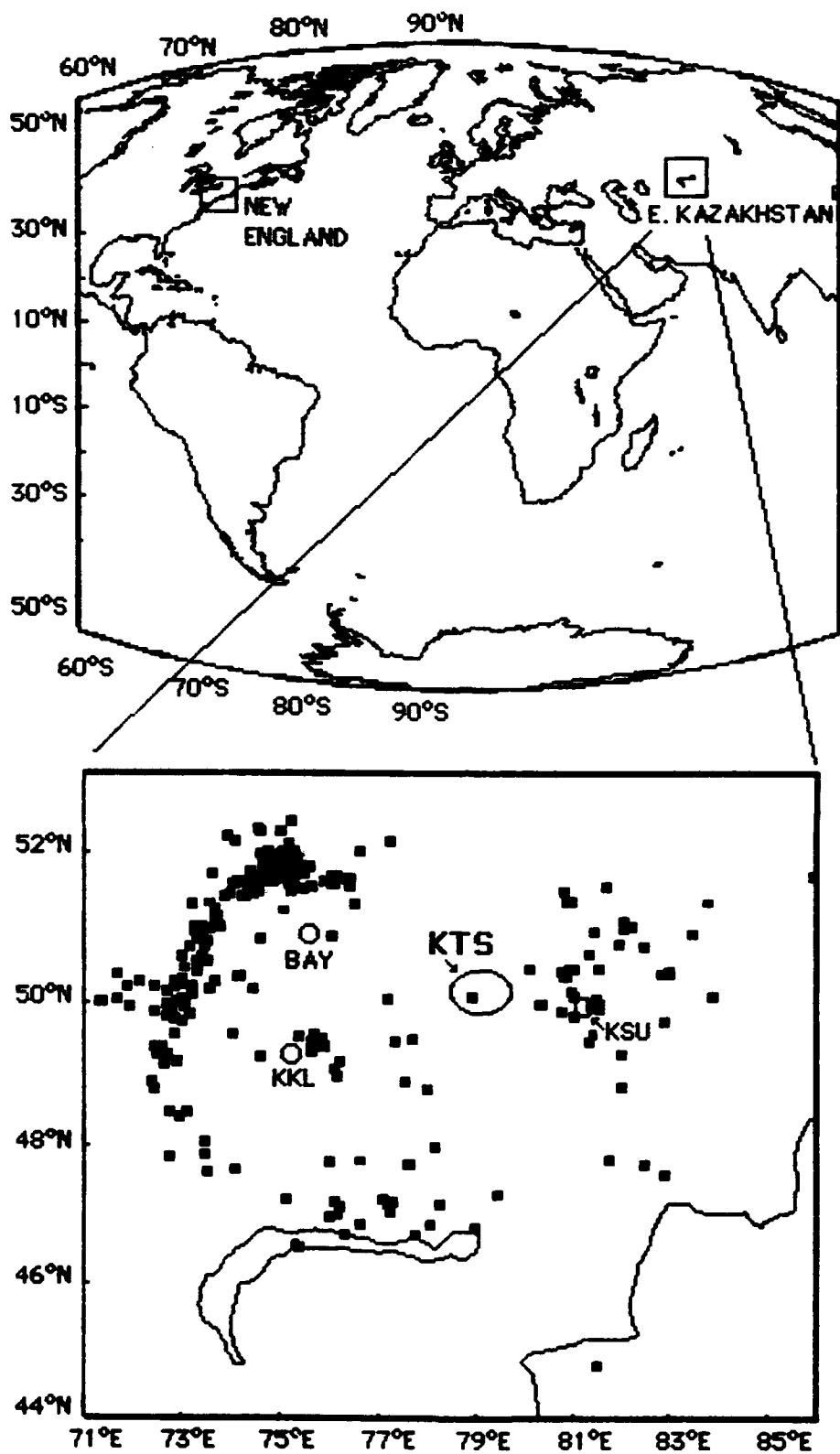
Figure 1: World map showing New England and Eastern Kazakhstan, and map of area surrounding the Kazakhstan Test Site (KTS). Open circles are stations operated by the NRDC in 1987, and closed squares are epicenters of seismic events recorded by those stations. The NRDC stations were located at Karkaralinsk (KKL), Bayanaul (BAY) and Karasu (KSU).

recorded ground motion are controlled not only by the source, but also by the characteristics of the propagation path of the seismic waves. Thus, effective seismic monitoring demands an ability to estimate source properties of an event based on a complicated seismic signal recorded at a distance. This is a particularly difficult problem at regional distances from the seismic sources because of the complicated path effects associated with seismic wave propagation in the Earth at regional distances (e.g. Pomeroy et al., 1982). With the availability of seismic data recorded in the vicinity of the KTS, we are now able to study seismic waves in that region to investigate the extent to which the path effects there differ from path effects in New England.

Figure 2 shows examples of seismograms recorded by the NESN from three quarry blasts and one earthquake in southern New England. All of the seismic events and propagation paths associated with the NESN seismograms discussed in this report are confined to geological features that are part of the Appalachian mountain belt. Rg is often observed from quarry blasts in New England and elsewhere; but unless an earthquake is very shallow, no Rg wave will be generated. For example, the seismogram shown in Figure 2(d) is recorded from a magnitude 2.9 (m_{bLg}) aftershock of the 1985 Ardsley, NY earthquake (m_{bLg} 3.8). The main event as well as its foreshocks and aftershocks all occurred at depths ranging from about 4.5 to 6.0 km (Seeber and Dawers, 1989). As would be expected for events at that depth, no Rg signals were recorded at the NESN stations from the Ardsley, NY earthquakes.

The configuration of stations installed by the NRDC is shown in the lower map in Figure 1. During 1987, the NRDC operated three stations in that region: Karakaralinsk (KKL), Bayanaul (BAY), and Karasu (KSU). In this study, I analyzed seismograms of quarry and mine blasts recorded by the NRDC stations. The Soviet events analyzed in this study, as well as the propagation paths associated with those events, are all contained within the Kazakhstan Accretionary Continent and the Central Asiatic Orogenic Belt described by Zonenshain et al. (1991). Since the region shown in that map is essentially aseismic, nearly all of the events recorded by the NRDC stations (indicated by closed squares in Figure 1) are "artificial" events, such as quarry and mining blasts (H.K. Given, personal communication).

In some cases, Rg is not observed very far from the source because it is attenuated by an upper layer of low Q in the shallow crust. This is one reason

North Branford, CT Quarry Blast

Station MD1 -- 36 km

Reed Gap, CT Quarry Blast

Station UCT -- 57 km

Reed Gap, CT Quarry Blast

Station NSC -- 74 km

Ardsley, NY Earthquake
Aftershock(10/21/85)

Station BCT -- 64 km
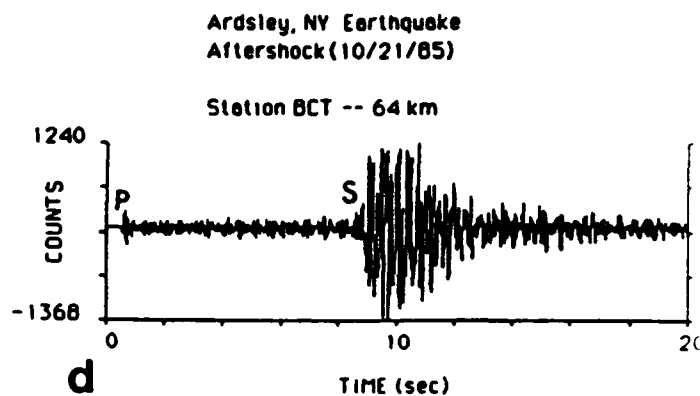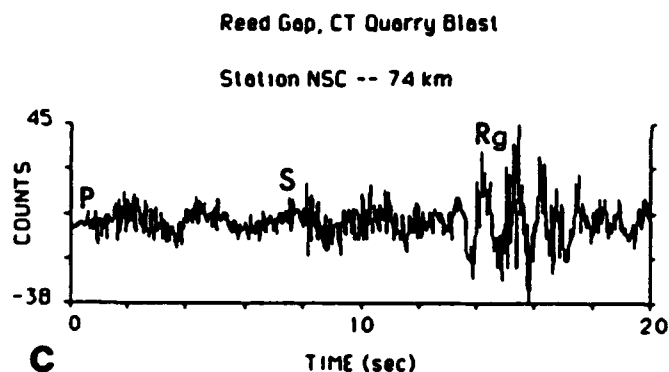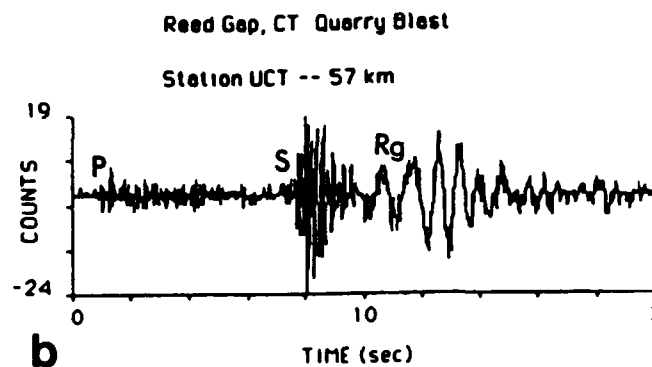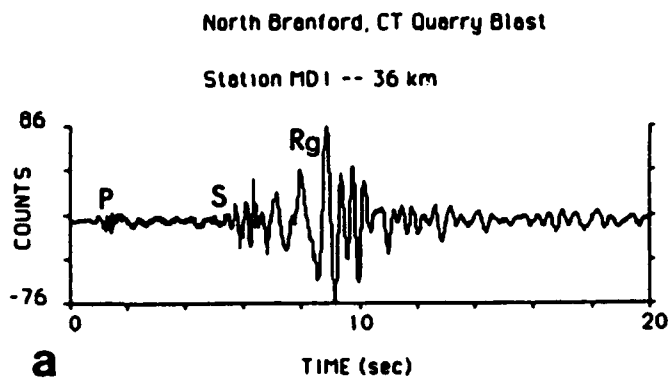
Figure 2: Seismograms of three quarry blasts and one earthquake recorded by the NESN in southern New England showing examples of Rg waves. The three examples of quarry blasts have prominent Rg waves because the sources are located near the Earth's surface, but the 5.2 km deep Ardsley, NY earthquake did not generate Rg waves.

why it is sometimes thought that Rg is not very practical as a depth discriminant. On the other hand, there are cases where strong Rg signals are observed at fairly large distances. The seismograms shown in Figure 3 are examples of the NRDC data analyzed in this study. Those seismograms were recorded at station KKL from a blast in the Karaganda mining area, located at a distance of 194 km from the station. The upper seismogram in Figure 3, which shows a very clear Rg signal, was recorded by the vertical component borehole instrument at KKL, and the seismogram directly below the borehole seismogram is the vertical component recorded at the surface. At KKL, the borehole instrument was located at a depth of 66 m.

The surface recording of the Karaganda blast shown in Figure 3 does not exhibit as prominent an Rg signal as the borehole seismogram. This could be due to a number of factors, including differences in recording characteristics for a seismometer buried at 66 m versus one installed at the surface. Differences in instrument response of the borehole and surface seismometers could, of course, be another factor affecting the strength of the Rg signals relative to other recorded waves. To evaluate the extent to which differences in instrument response are responsible for the difference in the quality of the Rg signal, the borehole seismogram was corrected for instrument response, and the resulting ground motion was convolved through the (somewhat narrower band) response of the surface instrument. The result is shown in the third seismogram from the top in Figure 3. The surface seismometers have a natural frequency of about 0.8 Hz, and the borehole seismometers have a natural frequency of about 0.2 Hz. Both types of instrument have anti-aliasing filters with a corner frequency of 80 Hz [see Berger et al. (1988) for details of the seismographic equipment installed at the NRDC sites]. It appears from Figure 3 that at least part of the difference in the relative strength of the Rg signal is due to the broader band response of the borehole instrument rather than the borehole placement per se.

To compare the NRDC seismograms directly with the NESN seismograms, we also corrected the surface NRDC seismograms for instrument response and then convolved the resulting ground motion through the NESN instrument response. An example of applying this procedure is shown in the bottom seismogram in Figure 3. The simulated NESN seismograms are

101-KKL
194 km
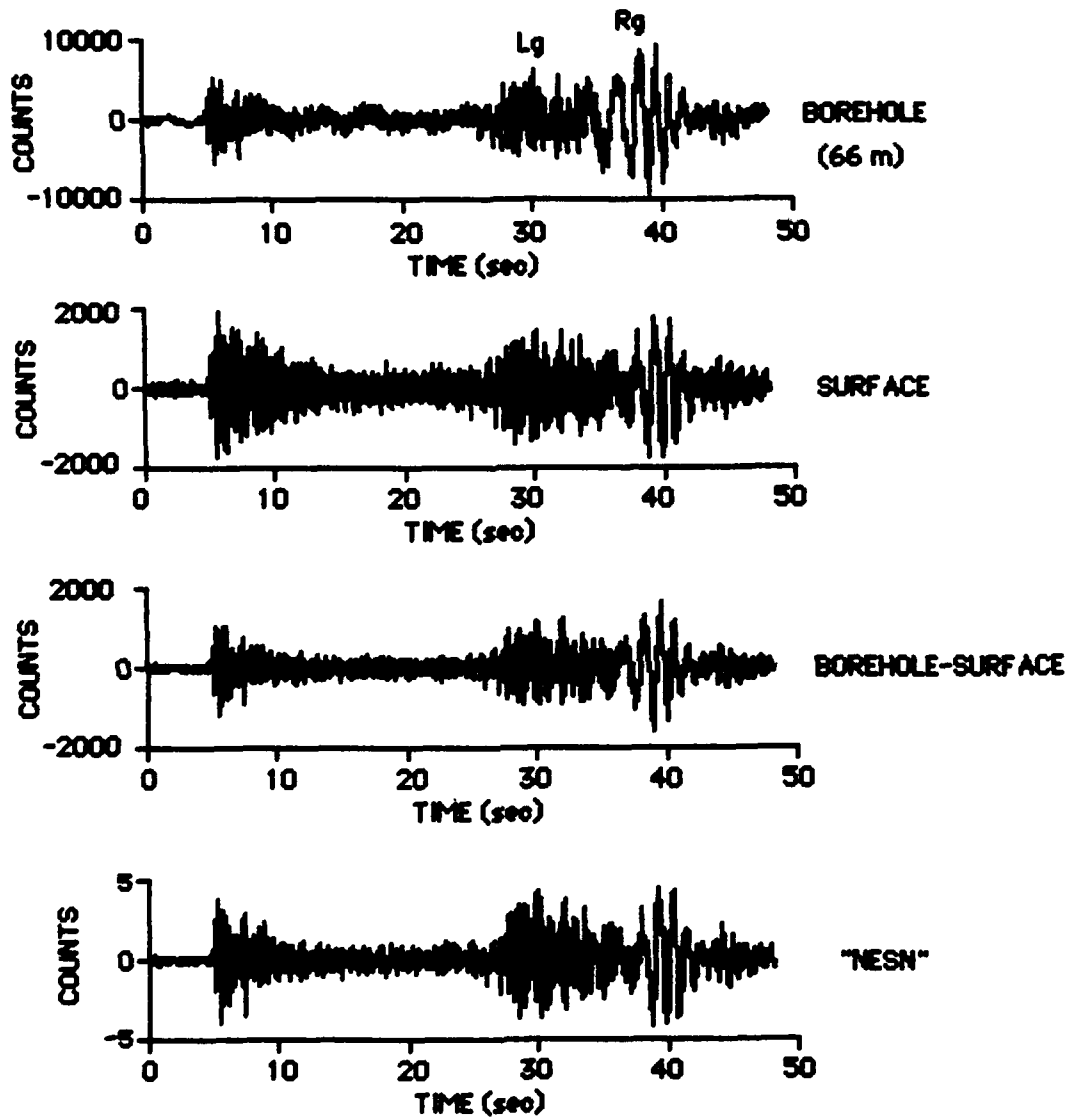
KARAGANDA BLAST

Figure 3: Example of a strong Rg signal recorded by borehole and surface instruments located at station KKL, a distance of 194 km from a blast in the Karaganda mining area.

used to investigate eastern Kazakhstan blasts as they would have been recorded if NESN type instruments were operating in that area.

The seismograms shown in Figure 4 illustrate that Rg waves can be observed at distances in excess of 250 km from blasts in the Karaganda mining area of eastern Kazakhstan. The processing and notation associated with Figure 4 is identical to that of Figure 3, but in this case the borehole instrument is located at a depth of 99 m, and the station (BAY) is 264 km from the blast. Based on the seismograms shown in Figure 4, it again appears that the broader band response of the borehole instruments was a significant part of the reason why stronger Rg signals are observed.

Although some combinations of strong source and efficient path yield Rg signals that propagate to significant distances, Figure 5 illustrates that some paths do not appear to be efficient propagators of Rg. The upper two seismograms are recorded from a fairly large blast at the Karagayly quarry (27 km from KKL and 172 km from BAY). The lower two seismograms in Figure 5 are from a smaller Karagayly blast. In both cases, a clear Rg signal is observed at KKL, but very little (if any) Rg energy is observed at BAY. A possible explanation of the difference in the strength of the Rg signal at the two stations is that there might be a strong radiation pattern for the Karagayly blasts (the azimuths of the two stations differ by about 90 to 100 deg). If, on the other hand, the lack of any observable Rg waves at station BAY cannot be explained by a pronounced radiation pattern, then it seems that the path from the Karagayly quarry to BAY attenuates essentially all of the Rg signal, even when quite strong signals are recorded at a closer station. In either case, this observation illustrates that the lack of an observed Rg wave is not necessarily an indicator of a deep event. However, a clearly identified Rg signal does indicate that the event was shallow.

4. Rg AS A DEPTH DISCRIMINANT

This research is part of a larger effort at Weston Observatory to develop methods of using Rg as a depth discriminant for earthquakes and explosions in various parts of the world. Recent results of the PI's research on Rg as a depth discriminant in New England are discussed in detail in Kafka (1990). This section presents a summary of those results.
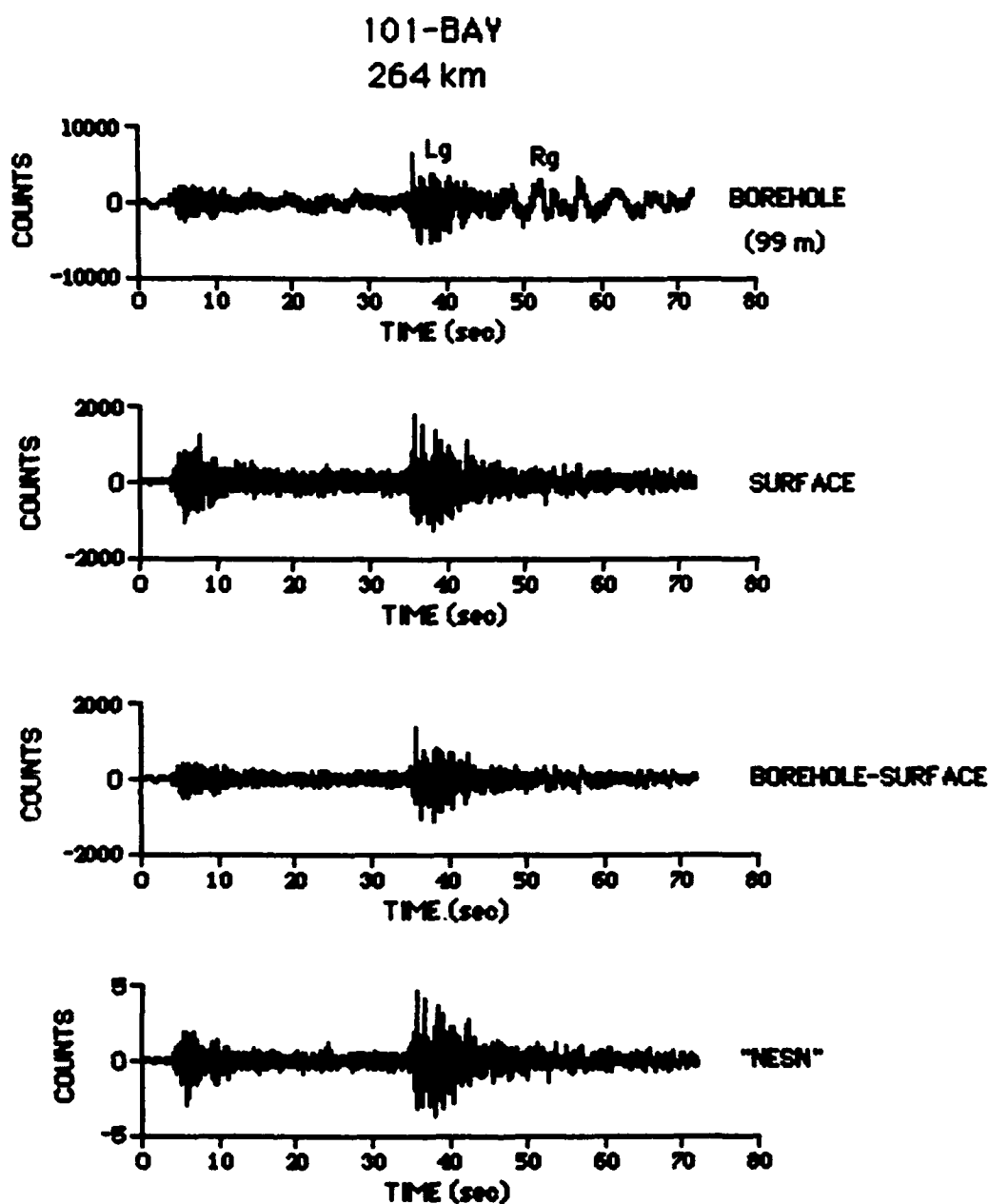
Figure 4: Seismograms recorded by borehole and surface instruments located at station BAY, a distance of 264 km from a blast in the Karaganda mining area.

Karagayly Quarry
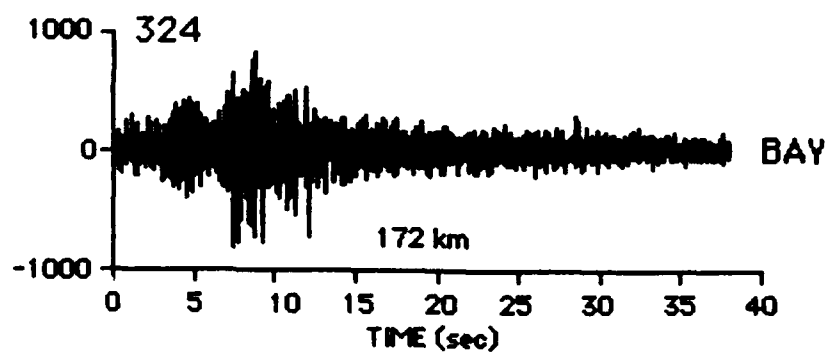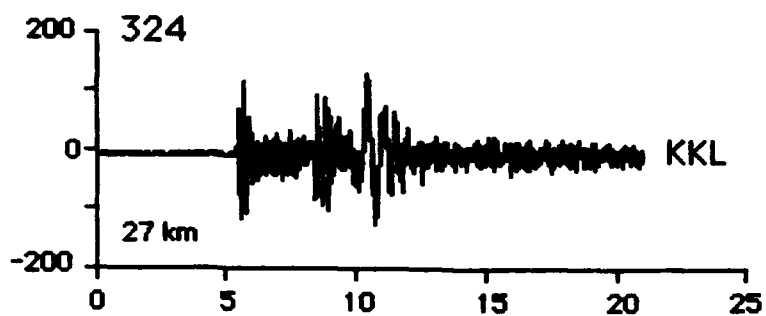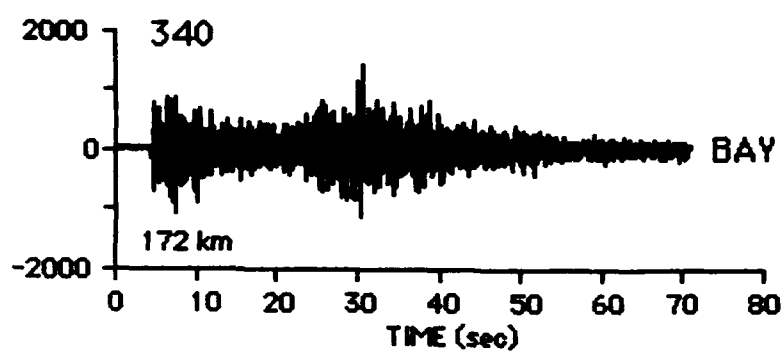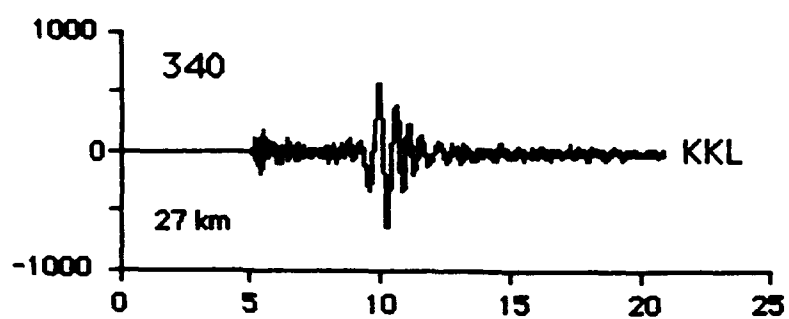
Figure 5: Seismograms recorded at stations KKL and BAY from blasts detonated at the Karagayly quarry.

Fundamental mode Rayleigh waves with periods ranging from about 0.4 to 2.5 sec (Rg) are often observed on seismograms of explosions and very shallow earthquakes in New England as well as in other parts of the world. Since Rg is a fundamental mode surface wave, it is not surprising that near-surface and very shallow-focus sources generate strong Rg signals. Bath (1975) proposed that Rg could be used as a depth discriminant for events recorded at regional distances. The principles underlying the use of Rg as a depth discriminant are that Rg amplitudes are very dependent on source depth, while amplitudes of other recorded phases such as Pg and Sg are, on average, less dependent on depth. Although these principles are straight forward, applying them to actual data can be problematic because of the complexity of short-period local and regional seismograms.

The strongest Rg signals recorded by the NESN are generally in the period range of about 0.5 to 1.5 sec. In that period range, Rg displacement is essentially confined to depths of about 1 to 5 km, with most of the Rayleigh-wave energy in the upper 2 or 3 km. Thus, sources deeper than about 4 km would not be expected to generate strong Rg signals. The task of developing a method for using Rg as a depth discriminant is therefore (to a large extent) a matter of developing a method for identifying the Rg phase and distinguishing it from other phases.

As with other regional discriminants, there are a number of reasons why using Rg as a depth discriminant is not as simple as the underlying principles suggest. For example (as mentioned above), Rg waves are likely to experience significant attenuation resulting from low Q material in the shallow crust. In addition, the radiation patterns of Rg are likely to be asymmetric for earthquakes (and possibly also asymmetric for some quarry blasts). Thus, there are a number of factors other than source depth that could be responsible for the lack of observed Rg waves. In spite of these problems, our work at Weston Observatory has shown that, at least in some areas of New England, Rg waves do have practical value as a depth discriminant. In practical situations, Rg might be used as one of several regional discriminants, none of which are likely to be without their own shortcomings.

My method of identifying Rg signals on seismograms involves estimating Rg/Lg ratios by using a narrow bandpass filter (NBF) analysis. This method is based on measuring amplitudes at particular periods and arrival times.

Figure 6(a) shows the analysis of a seismogram with a prominent Rg wave as well as a prominent arrival in the S and Lg time windows. Since S and Lg are difficult to separate at these distances and frequencies, the notation "Lg" will (in the remainder of this report) refer to the entire wave train from the onset of the S wave to the end of the S and Lg coda. Hence, S and Lg are analyzed as one (complicated) wave train.

The seismogram in Figure 6(a) has high amplitudes in both the Lg and Rg arrival time-frequency windows. The situation is different for the seismogram in Figure 6(b), where Rg has much larger amplitudes than Lg.

The seismogram of the Erving, MA earthquake (June 14, 1984) recorded at station WES has a small (but observable) Rg wave, and an NBF analysis of that seismogram is shown in Figure 7(a). Relatively high amplitudes are observed in both the Lg and Rg parts of the velocity-period plane of the Erving, MA - WES seismogram.

These examples suggest that the ratios of Rg amplitudes to Lg amplitudes can be estimated by narrow bandpass filtering each seismogram and then taking the ratios of amplitudes within the appropriate velocity-period windows. To do this it is necessary to estimate group velocity dispersion of Rg waves in the area being investigated. Figure 7(b) shows the results of a number of studies of Rg dispersion in New England.

Figure 7(c) illustrates the method of estimating Rg/Lg ratios. The amplitudes are entered into a grid, with a given cell representing the amplitude at a particular point on the velocity-period plane. All amplitudes in the cells labelled Lg are averaged to give an estimate of the Lg amplitude. Then the amplitudes in the cells labelled Rg are averaged to give an estimate of the Rg amplitude. The Rg/Lg ratio is then calculated by dividing the "Rg" average by the "Lg" average.

Figure 8 shows histograms of Rg/Lg ratios for blasts and earthquakes recorded in New England. Two earthquake sequences for which good depth estimates were available from aftershock surveys were the 1985 Ardsley, NY earthquakes (which occurred at a depth of about 5.2 km) and the 1987 Moodus, CT earthquakes (which occurred at a depth of about 1.6 km). Additional information about these earthquakes is given in Kafka (1990). There is a clear distinction between the Rg/Lg ratios for the blasts and for the Ardsley, NY earthquakes, which would be expected because of the 5.2 km depth. Also shown in Figure 8 are the ratios for the Moodus, CT earthquake
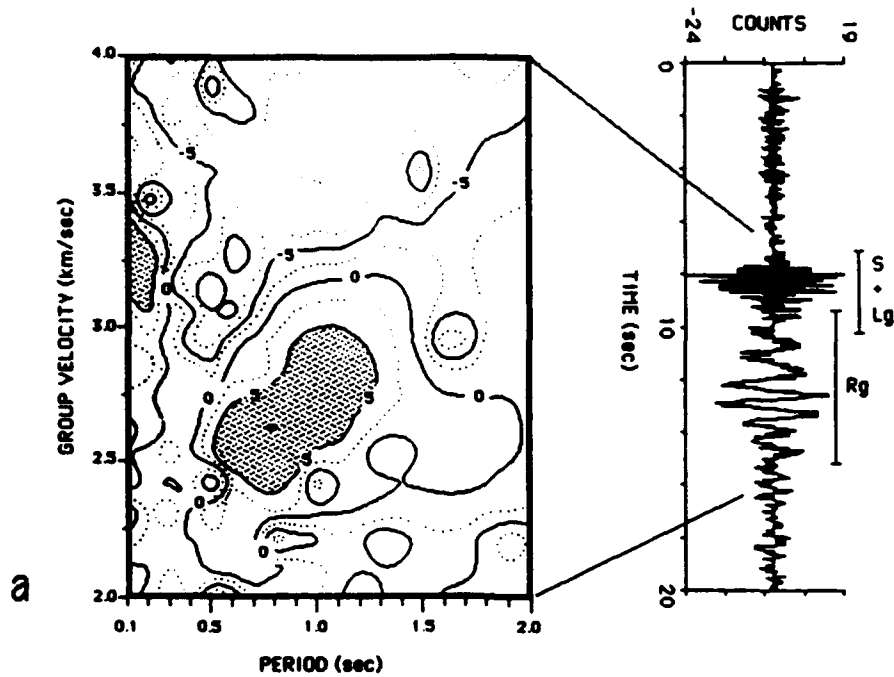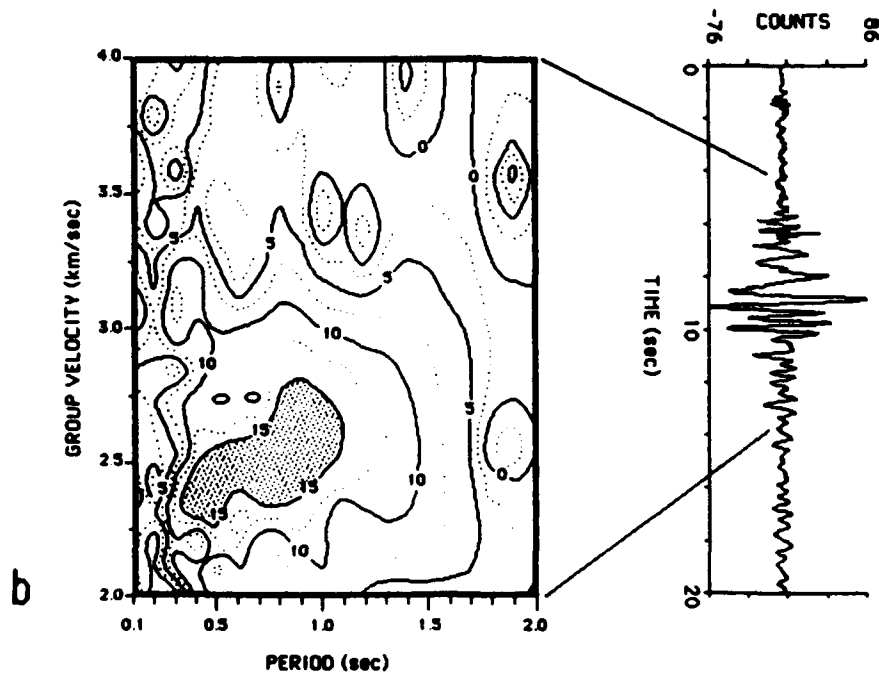
Figure 6: Narrow bandpass filter analysis of (a) seismogram recorded at station UCT from Reed Gap, CT quarry blast, and (b) seismogram recorded at station MD1 from North Branford, CT quarry blast. The vertical axis in these contour plots is amplitude, and each unit of amplitude represents a difference of 2 db. Shaded areas indicate highest amplitudes.
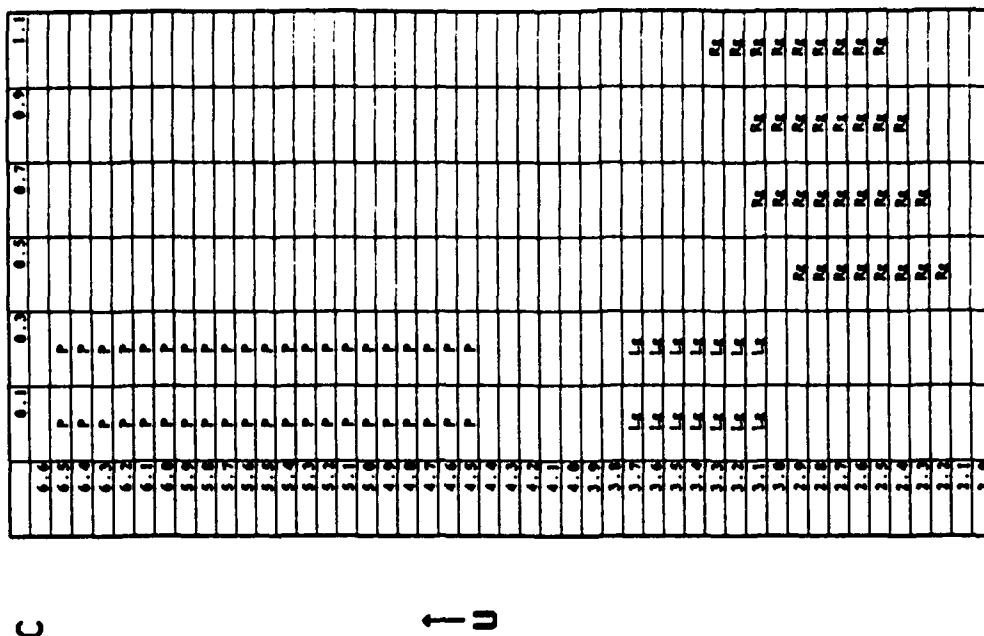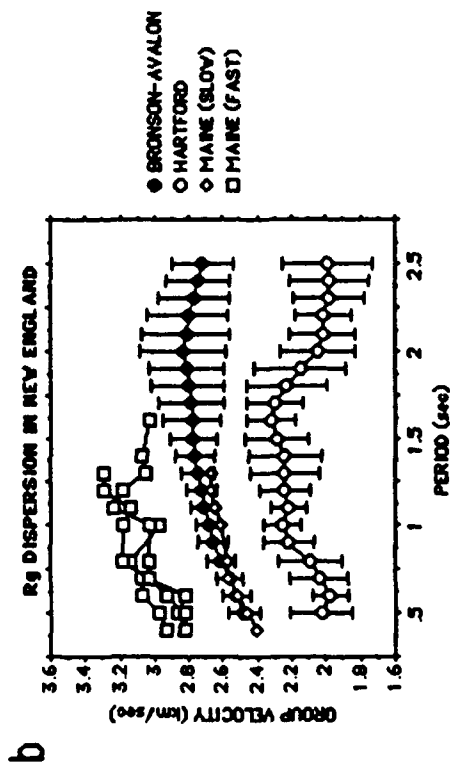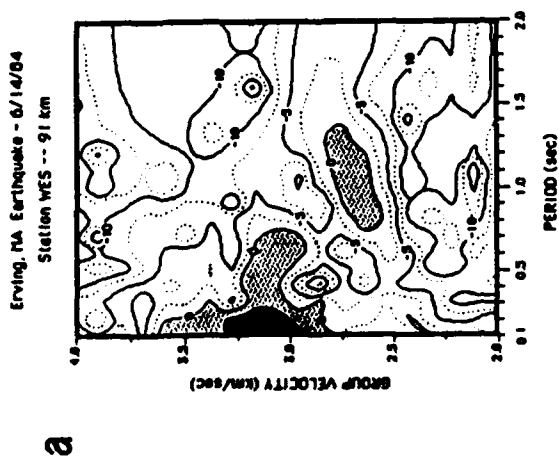
Figure 7: (a) Narrow bandpass filter analysis of seismogram recorded at station WES from the Erving, MA earthquake. (b) Summary of Rg group velocity dispersion in New England, from Kafka and Reiter (1987) and Kafka and Bowers (1991). (c) Method of calculating Rg/Lg ratios from narrow bandpass filter analysis. U is group velocity in km/sec, and T is period in sec.

Rg/Lg Amplitude Ratios and Focal Depth
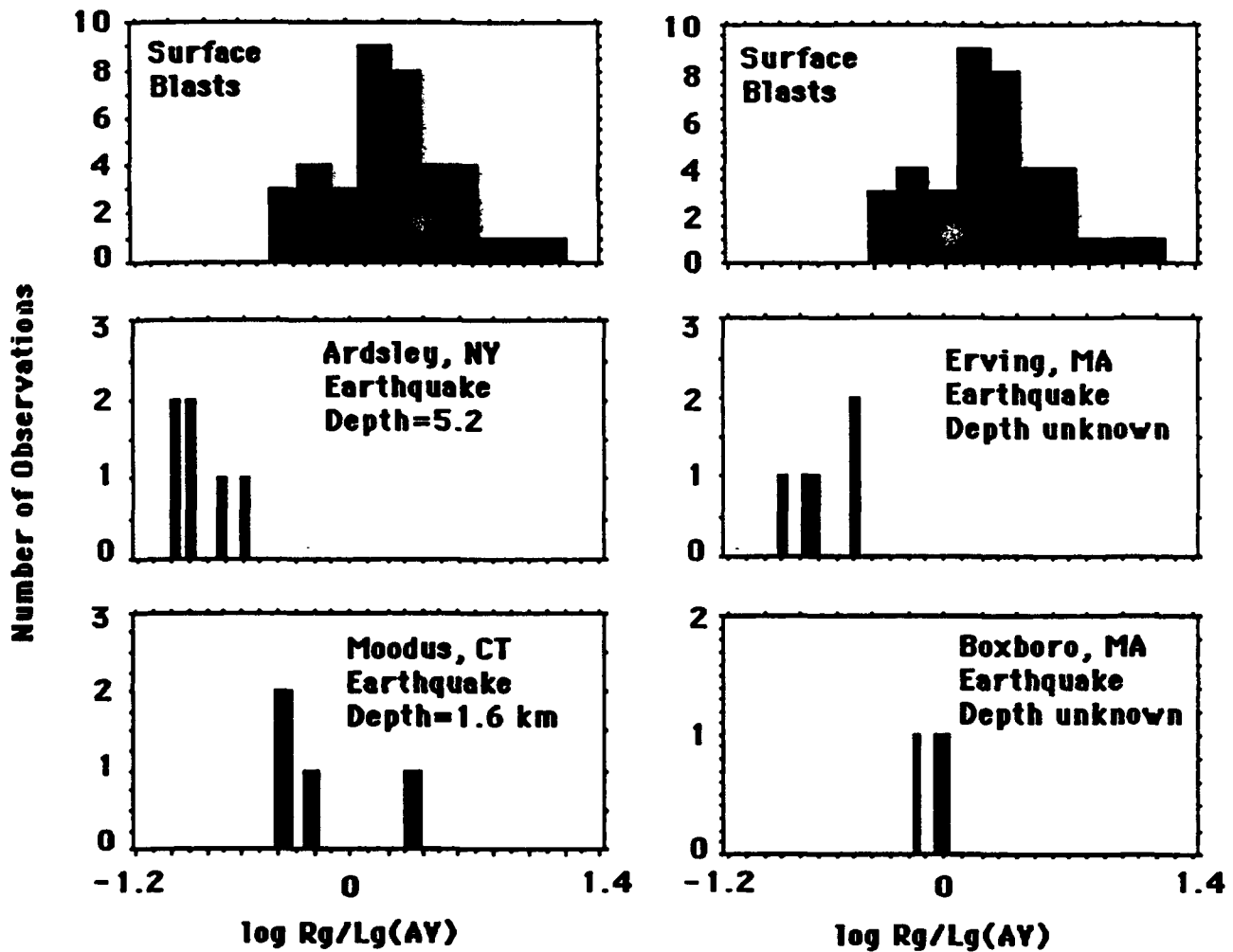
Figure 8: Histograms of Rg/Lg ratios for surface blasts and for four earthquakes.

and for two other earthquakes that had poorly constrained depths. All of the ratios for the Moodus earthquake overlap with those of the blasts, which would be expected because that earthquake is known to be shallow. The average value of Rg/Lg for the Moodus earthquake (0.93) is lower than the average value for the blasts (2.50), which is consistent with the 1.6 km depth of the earthquake.

Based on the average Rg/Lg ratios for the Erving, MA and Boxboro, MA earthquakes, I estimate their depths to be about 3.0 km and 1.2 km, respectively. The details of how I obtained these depth estimates are given in Kafka (1990).

5. ANALYSIS OF THE NRDC SEISMOGRAMS AND COMPARISON OF NEW ENGLAND AND EASTERN KAZAKHSTAN

The primary goal of this study was to compare Rg signals recorded in New England with Rg signals recorded in eastern Kazakhstan. The first step in this analysis was to compile a list of known quarry and mining blasts in the vicinity of the NRDC stations (Table 1). Although there are a very large number of suspected blasts in both the NRDC and the NESN data sets, it was important to base this work on seismograms of confirmed blasts in both areas. This assured me that I was in fact analyzing seismograms from events with known depth, and that at least one of the source parameters was essentially the same (i.e. zero km depth) for all events used for this comparison.

Table 1 was compiled from two sources: (1) a list of blasts provided by H.K. Given, who participated in the field expedition to install the NRDC instruments, and (2) information published by Thurber et al. (1989) on quarry and mining blasts in the vicinity of the KTS. A similar list of confirmed quarry blasts that were detonated during 1989 at the San-Vel quarry in Littleton, MA was compiled from information provided by the operators of that quarry.

One of the most difficult tasks of this study turned out to be the transfer of the NRDC seismograms from the Center for Seismic Studies (CSS) to the Boston College VAX computer (BC-VAX). The procedure recommended by personnel at the CSS was to copy the data to computer tapes at the CSS and then to have

Table 1

CONFIRMED QUARRY AND MINING BLASTS
RECORDED BY THE NRDC STATIONS

Event Time (day:hr:min:sec)	Event Number	Location	Latitude, Longitude	Stations
100:09:52	30*	Karagayly	49.43, 75.82	KKL
113:09:51	44*	Karagayly	49.48, 75.83	KKL
115:08:25	58*	Karagayly	49.42, 75.86	KKL
117:09:47	71*	Karagayly	49.51, 75.38	KKL
118:08:47	81*	Karagayly	49.44, 75.64	KKL
134:09:36:16.4	155*	Karaganda	50.192, 74.157	BAY, KKL
135:09:08:35.2	98*	Ekibastuz	51.709, 75.514	BAY, KKL
135:10:35:00.3	101*	Karaganda	49.304, 72.712	BAY, KKL
140:09:54	318*	Karagayly	49.37, 75.89	BAY, KKL
141:08:58	324*	Karagayly	49.39, 75.88	BAY, KKL
141:09:16:43.3	326	Karaganda	50.744, 73.279	BAY, KKL
143:08:49:22.7	340*	Karagayly	49.275, 75.738	BAY, KKL
145:09:26:43.7	350*	Ekibastuz	51.670, 75.454	BAY, KKL
145:09:56:40.9	353*	Ekibastuz	51.743, 75.316	BAY, KKL
146:05:31:04.8	357*	Ekibastuz	51.819, 74.797	BAY, KKL
146:08:33:26.5	359	Ekibastuz	51.760, 75.571	BAY, KKL
162:12:41:04.0	378	Ekibastuz	51.454, 75.488	BAY, KKL
162:12:50:34.3	498	Ekibastuz	51.677, 75.525	BAY, KKL
239:08:52:53.0	839	Ekibastuz	51.213, 74.302	BAY, KKL
239:09:37:34.8	845	Balkash	46.900, 77.389	BAY, KKL, KSU
244:09:08:52.0	967	Balkash	46.924, 77.241	BAY, KKL, KSU
245:08:02:10.2	993	Ekibastuz	51.639, 75.481	BAY, KKL, KSU
245:08:53	997*	Karagayly	49.41, 75.66	BAY, KKL
248:08:47	1081	Karagayly	49.41, 75.82	

* on first CSS tape

Events shown with origin times in seconds and tenths of seconds are from Thurber et al. (1989); other events are from a list provided by H.K. Given (personal communication). Locations shown with latitude and longitude in thousandths of a degree are from Thurber et al. (1989); those shown with latitude and longitude in hundredths of a degree are from the CSS data files.

the investigator read the tapes onto the computer at the investigator's institution. This caused a number of problems because the different computers and operating systems did not read the data in the same way. The bulk of the NRDC data for this study was transferred to Boston College on two tapes that, after a significant effort, have now been successfully read onto the BC-VAX. Part of the problem was that after we successfully read the first tape, the CSS personnel changed the way the data was stored when it was written on the second tape. A third tape is now being prepared by the CSS. Unfortunately, because the tape format was changed after we had finally succeeded in reading the first tape, we encountered additional delays in reading the second tape. The results discussed in this report are, therefore, primarily based on data that was on the first tape.

Seismograms stored on the first tape were analyzed to obtain estimates of Rg/Lg ratios for a sample of events in eastern Kazakhstan. That tape was created at the CSS by searching for all NRDC seismograms that had Rg signals (identified by CSS analysts) that were greater than 100 nanometers of ground displacement (R. Baumstark, personal communication). I searched through the events on that tape to determine which of the events listed in Table 1 had seismic waveforms that could be analyzed using the methods described above.

To obtain accurate estimates of Rg/Lg ratios for the NRDC seismograms, it is necessary to estimate the Rg group velocity dispersion in the area surrounding the KTS. Figure 9 shows a comparison of Rg dispersion results for New England with results for seven paths in eastern Kazakhstan. For the Soviet dispersion curves, I used the origin times and locations given by Thurber et al. (1989). The source parameters of these Soviet blasts were determined from a sparse network of only two or three stations. Thus, locations of these events are only accurate to within at best a few km, and in some cases, location errors of these blasts could be greater than 10 km (Thurber et al., 1989). It is therefore difficult to evaluate the details of how the dispersion results for eastern Kazakhstan are related to geological structures in that region. Nonetheless, it does appear that the range of group velocities is about the same for the two regions, and that the methods developed for New England can, as a first approximation, be applied directly to the NRDC seismograms.

Figure 10 shows an example of a comparison between a New England quarry blast and an eastern Kazakhstan quarry blast recorded at about the
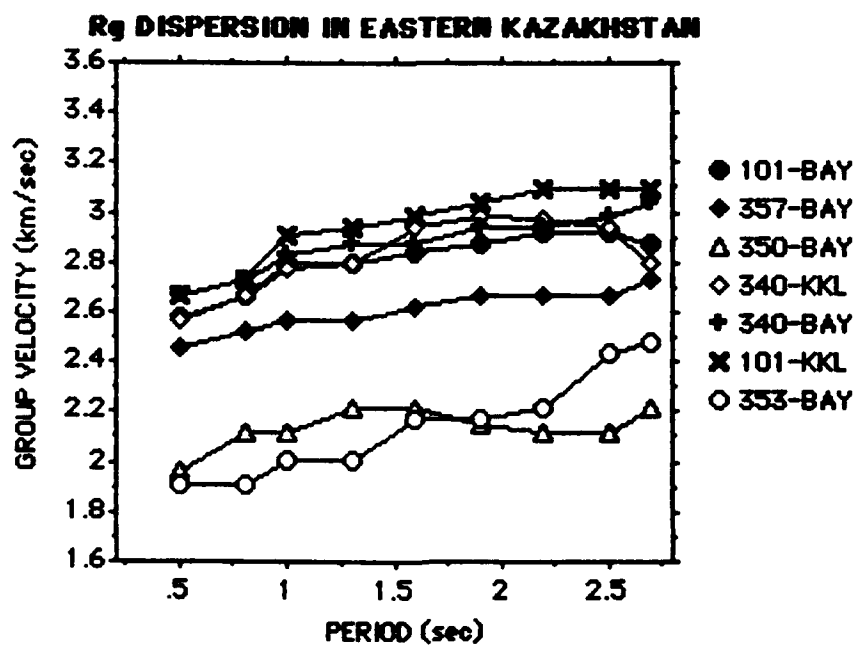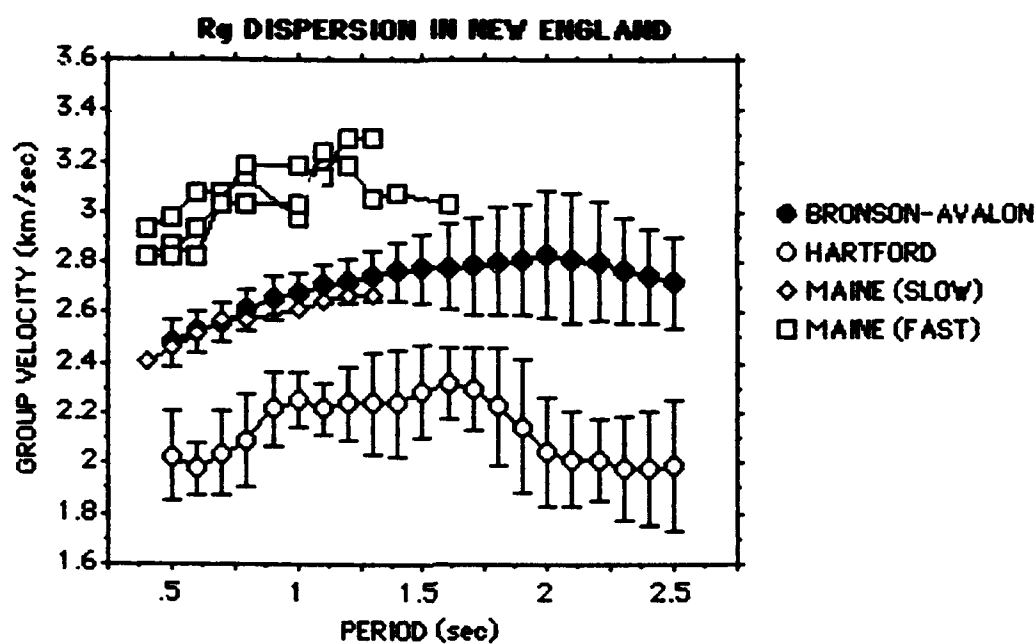
Figure 9: Comparison of Rg dispersion in New England and eastern Kazakhstan.
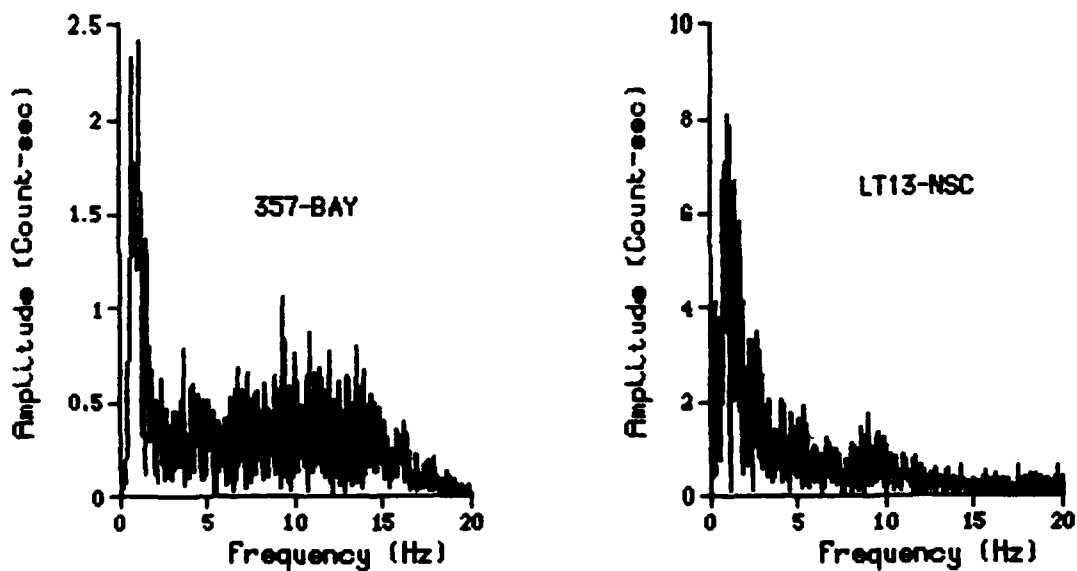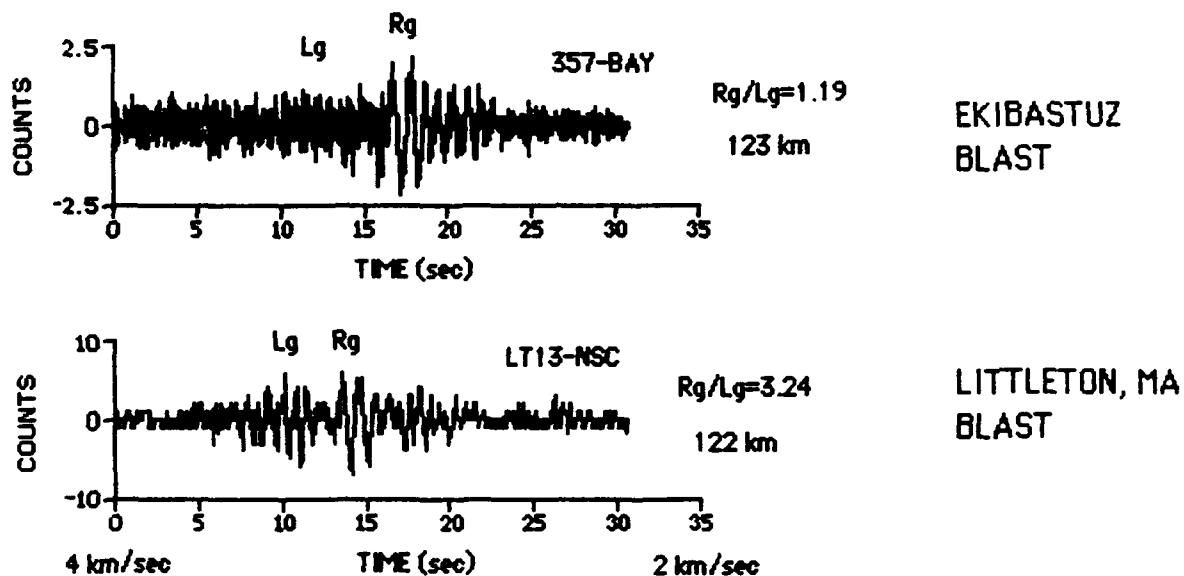
Figure 10: Example of a comparison between a New England quarry blast and an eastern Kazakhstan quarry blast recorded at about the same distance. The upper seismogram is an NRDC seismogram (recorded by a surface seismometer) processed to simulate an NESN instrument response (as described in the text).

same distance. The upper seismogram in Figure 10 is an NRDC seismogram (recorded by a surface seismometer) processed to simulate an NESN instrument response (as described above). The NESN stations have a 1 Hz seismometer and a low-pass filter with a corner frequency of 12.5 Hz. The NRDC instruments have a broader band response than the NESN instruments (0.8 Hz seismometer and 80 Hz low pass filter for the surface instruments). Thus, the method of processing the NRDC seismograms to simulate NESN seismograms essentially involves applying a narrower bandpass filter than that of the original instrument. The resulting seismograms are, therefore, not likely to contain frequency components that are artifacts of the data processing method.

The NRDC seismogram shown in Figure 10 was recorded at station BAY, a distance of 123 km from a blast at Ekibastuz. The NESN seismogram in that Figure was recorded at station NSC, a distance of 122 km from a quarry blast at the San-Vel quarry in Littleton, MA. The Soviet blast has greater signal strength at higher frequencies, even after correcting for the narrower band response of the NESN stations. The Rg/Lg ratio was calculated for both seismograms in Figure 10 following the procedure outlined above in Section 4. For the Ekibastuz blast the Rg/Lg ratio was found to be 1.19, and for the Littleton quarry blast the Rg/Lg ratio was 3.24.

In order to make as direct a comparison as possible between Rg/Lg ratios in eastern Kazakhstan and New England, I decided to compare NRDC seismograms with NESN seismograms using the following criteria in both cases: (1) strong seismic signals repeatedly recorded at the same station from blasts in the same source region, (2) similar path lengths for the two different cases, and (3) evidence of similar Rg dispersion. The best candidates that I was able to find for this comparison were seismograms from the Karagayly quarry recorded at KKL, and seismograms from the Littleton quarry recorded at WES. The path from Littleton to WES is contained within a region that has Rg dispersion similar to that shown in Figure 9 for the Karagayly-KKL path, and the path lengths were 25 km and 27 km, for New England and eastern Kazakhstan, respectively. From the list of Littleton quarry blasts, I found seventeen blasts that triggerred the NESN and were recorded at station WES. Out of ten Karagayly blasts listed in Table 1, all were on the first CSS tape except for one (event 1081).

The Karagayly-KKL seismograms were processed to simulate the NESN instrument response. Rg/Lg ratios were then estimated for the Karagayly-KKL seismograms and the Littleton-WES seismograms using the method described above. The similarities and differences between the two sets of Rg/Lg ratios is illustrated in Figures 11 and 12. This comparison s, of course, only based on a small number of seismograms, and it is not possible to make any general conclusions about differences between Rg/Lg ratios in the two regions. Nonetheless, it is worth noting that there is significant overlap between the two samples of ratios and that, on average, the Karagayly-KKL seismograms have slightly higher ratios than the Littleton-WES seismograms. Figure 11 shows seismograms corresponding to the highest and lowest Karagayly-KKL ratios (17.54 and 1.91, respectively) as well as the highest and lowest Littleton-WES ratios (4.21 and 0.62, respectively). Figure 12 shows histograms of the results for Rg/Lg ratios of the Karagayly-KKL and Littleton-WES seismograms.

Since the NRDC seismograms were processed to simulate the NESN instruments, the differences in Rg/Lg ratios for the Kargayly-KKL and Littleton-WES seismograms are apparently caused by something other than differences in instrument response. These differences in the ratios could be caused by differences in the source, the path, and/or the site. It is, of course, difficult to separate out those effects based only on the small number of seismograms that were analyzed in this study. Nonetheless, the comparison between the Karagayly-KKL and the Littleton-WES seismograms was designed to be as direct a comparison as possible, and to motivate future studies of differences between Rg and Lg wave excitation and propagation in these two regions.

6. RECOMMENDATIONS

In spite of the problems associated with estimating depth from Rg/Lg ratios, it seems that, at least in so e cases, Rg waves do have practical value as a depth discriminant. To apply this research in practical situations, it will be necessary to evaluate the effects of the source, path, and site on Rg and Lg amplitudes for each particular situation. Separating these effects from each other is an issue for essentially any regional seismic discriminant that would
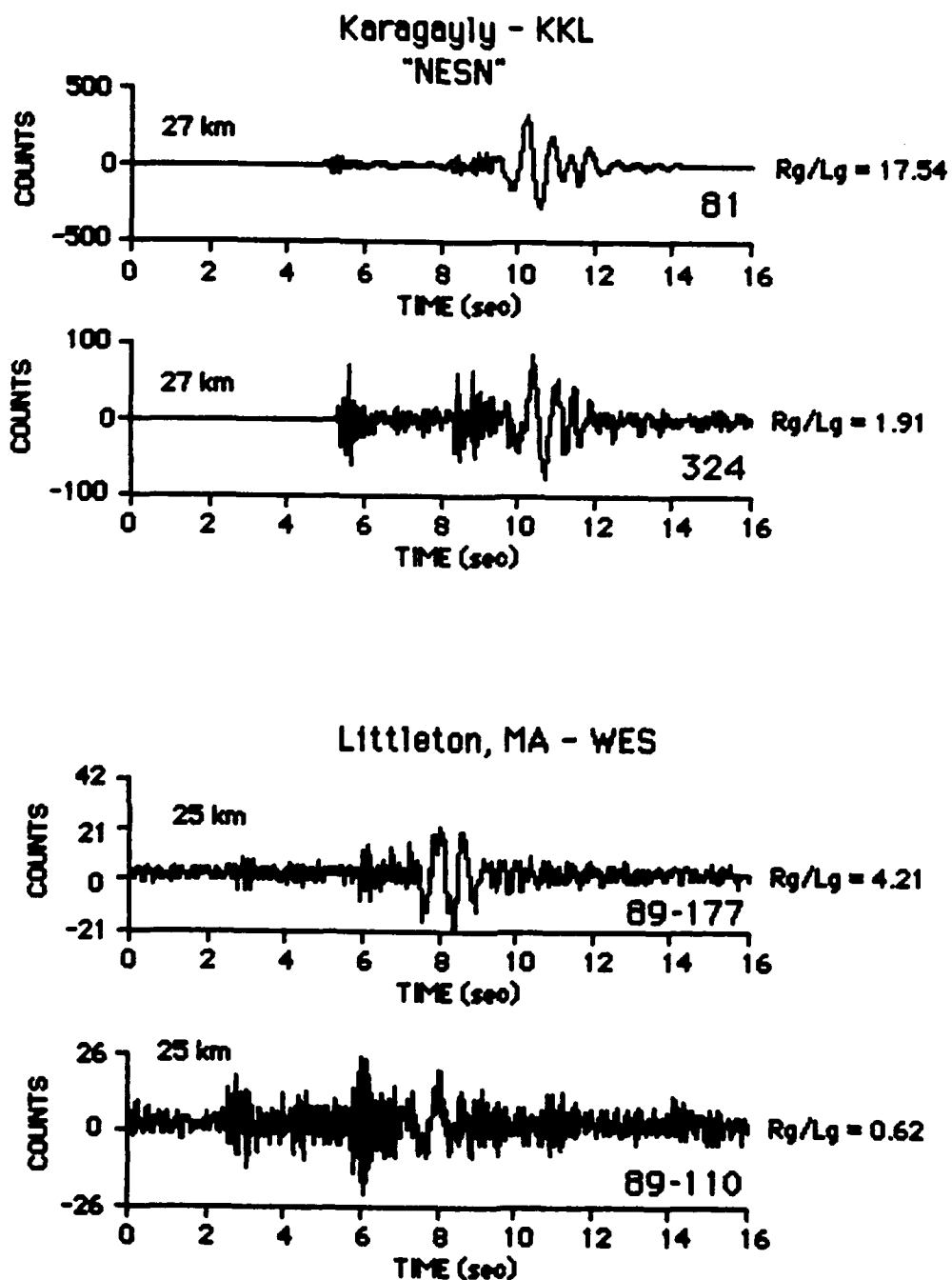
Figure 11: Seismograms corresponding to the highest and lowest Rg/Lg ratios for the Karagayly-KKL path and for the Littleton-WES path. The Karagayly-KKL seismograms were processed to simulate the NESN instrument response.
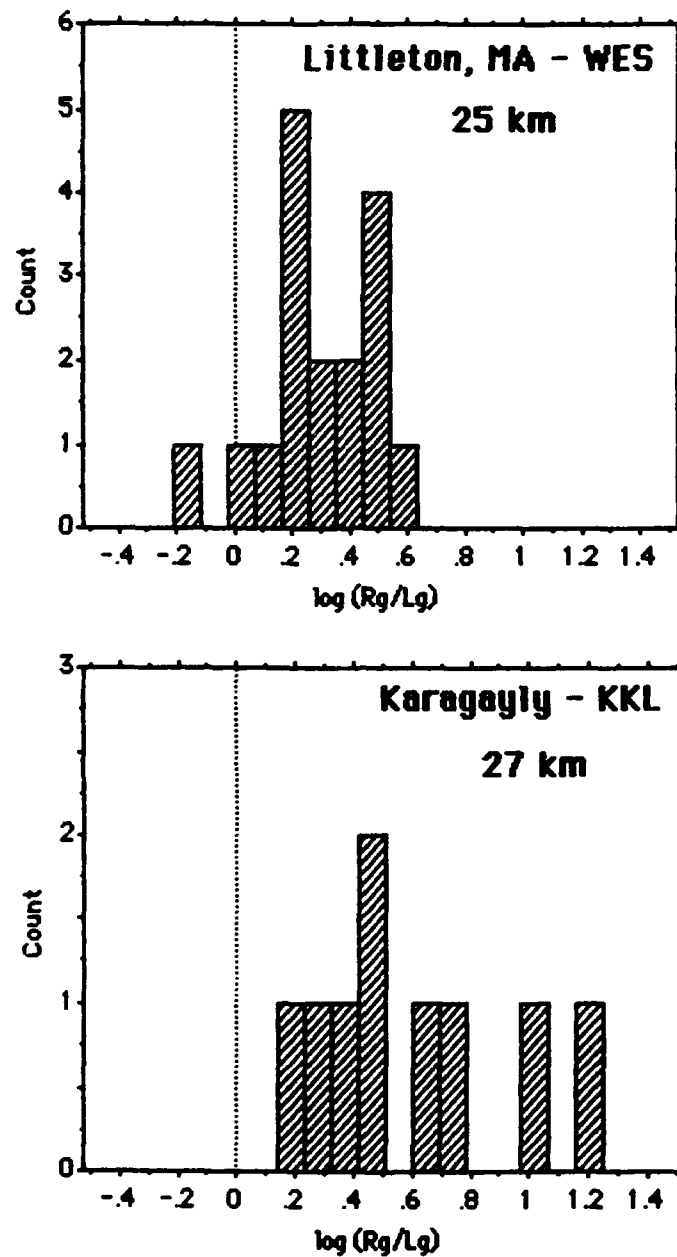
Figure 12: Histograms of Rg/Lg ratios of the Karagayly-KKL and Littleton-WES seismograms.

be used for monitoring small explosions. If Rg proves to be a useful discriminant in practical situations, it would have to be applied as one of several regional discriminants, none of which are likely to be without their own shortcomings.

The primary reason why it is difficult to make general conclusions about differences between Rg/Lg ratios in the two regions discussed in this report is that each blast analyzed here was well-recorded by at most three stations (and usually by only one or two stations). With such a sparse network of recording stations, it is of course difficult to design experiments to evaluate the various possible effects on Rg and Lg amplitudes.

To evaluate the practicality of applying the Rg/Lg ratio method discussed in this report, it would probably be best to use two complementary approaches. First, "controlled" amplitude experiments with a denser distribution of stations should be conducted in some area to evaluate the range of effects that the source, path, and site can have on Rg and Lg amplitudes propagating in a given geological region. Second, in the Soviet Union (or other foreign countries) where the U.S. would be allowed, at most, limited access, it is probably best to "calibrate" specific source-path-site-station combinations by analyzing many seismograms recorded at a given station from a given source region. Hopefully, this type of procedure would yield sufficient data that it would be possible to catalogue characteristic features of different types of events. It seems that the range of Rg/Lg ratios for a given type of event recorded at a given station should be considered as one of the possible identifying characteristics.

At Weston Observatory, we are in the process of conducting experiments in New England to investigate the range of effects that the source, path, and site can have on seismic wave amplitudes recorded at regional distances. Field stations are being installed at a range of azimuths and distances from quarry blasts detonated in New England. The overall question that we are attempting to answer with these experiments is: To what extent can observed differences in amplitudes of regional seismic phases be interpreted as evidence of the characteristics of the source? We are concentrating our efforts on investigating source, path and site effects on Rg and Lg amplitudes.

ACKNOWLEDGEMENTS

7. REFERENCES

Bath, M., Short-period Rayleigh waves from near-surface events, Phys. Earth Planet. Interiors, 10, 369-376, 1975.

Berger, J., H.K. Eissler, F.L. Vernon, I.L. Nersesov, M.B. Gokhberg, O.A. Stolyrov and N.T. Tarasov, Bull. Seis. Soc. Am., 78, 1744-1758, 1988.

Kafka, A.L. and A.K. Bowers, Corrections to Rg group velocity dispersion in southwestern New England, Seis. Res. Lett., in press, 1991.

Kafka, A.L., Rg as a depth discriminant for earthquakes and explosions: A case study in New England, Bull. Seis. Soc. Am., 80, 373-394, 1990.

Kafka, A.L. and E.C. Reiter, Dispersion of Rg waves in southeastern Maine: evidence for lateral anisotropy in the shallow crust, Bull. Seis. Soc. Am., 77, 925-941, 1987.

Office of Technology Assessment, Seismic Verification of Nuclear Testing Treaties, G. E. van der Vink (Project Director), U.S. Congress, OTA-ISC-361, U.S. Government Printing Office, Washington, D.C., May 1988.

Pomeroy, P.W., W.J. Best and T.V. McEvilly, Test ban treaty verification with regional data - a review, Bull. Seis. Soc. Am., 72, S89-S129, 1982.

Seeber, L. and N. Dawers, Characterization of an intraplate seismogenic fault in the Manhattan Prong, Westchester Co., N.Y., Seis. Res. Lett., 60(2), 71-78, 1989.

Simarski, L.T., Seismic glasnost opens U.S.S.R. to U.S. Researchers, EOS, Trans. Am. Geophys. Un., 72, 9-10, 1991.

Thurber, C, H. Given, and J. Berger, Regional seismic event location with a sparse network: Application to Eastern Kazakhstan, USSR, J. Geophys. Res., 94, 17767-17780, 1989.

Zonenshain, L.P., J. Verhoef, R.Macnab and H. Meyers, Magnetic imprints of continental accretion in the U.S.S.R., EOS, Trans. Am. Geophys. Un., 72, 305-310, 1991.

Time-of-Flight Simulations

of Collisions of

$H_2^{18}O^+$ with D_2O

by

C. Randal Lishawa, Ph.D.

Department of Physics and Engineering

Utica College of Syracuse University

Utica, NY

Acknowledgements

Time-of-Flight Simulations
of Collisions of
$H_2{}^{18}O^+$ with D_2O

ABSTRACT

Time-of-flight simulations have been conducted on the $H_2{}^{18}O^+/D_2O$ reaction system at a collision energy of 29.5 eV within a modified spectator stripping model. The time-of-flight spectra and calculated cross section for the reaction are compared to previously published data. It is possible to match the maximum in the time-of-flight experiment with the calculated results assuming the reaction occurs at a 1 Å center of mass separation although the shape of the distribution is not well reproduced. The calculated cross section is also poorly reproduce, being an order of magnitude too large.

Time-of-Flight Simulations

of Collisions of

$H_2{}^{18}O^+$ with D_2O

I. INTRODUCTION

The Spacecraft Contamination group at the Geophysics

Laboratory has been investigating the effects of the

contaminant cloud that accompanies low orbit spacecraft

(including the Space Shuttle). Ions present in the earth's

atmosphere will interact with the contaminant species

surrounding the spacecraft. With orbital velocities around

7.7 km/s (approximately 5.6 eV translational energy for

H_2O), many of the reactions occur at suprathermal (> 1 eV)

energies. Very few studies of chemical reactions have been

undertaken at these energies. Lishawa et.al.[1] have

recently measured the energy dependent reaction cross

section and the time-of-flight product distributions for the

reactions of H_2O^+/H_2O using isotopically labelled reactants

to differentiate between possible reaction channels.

Experiments designed to look at the internal energy

distributions of the products are currently underway.[2]

This report describes a computer simulation of the

translational energy distribution of a single reaction product channel ($H_2^{18}O^+$ + D_2O → $DH_2^{18}O^+$ + OD) resulting from the ion-molecule collisions of water vapor.

II. PROGRAM DESCRIPTION

The program uses a fourth-order Runge-Kutta algorithm[3] to compute a classical trajectory simulation for the reactions of suprathermal molecular ions colliding with stationary dipolar molecules. Predictor-corrector methods were not used because of the simplicity of the required derivatives. The classical potential energy function for an ion/polarizable-dipole system from electromagnetism is

$$V(R) = -\frac{\alpha q^2}{2R^4} - \frac{\mu q}{R^4}\cos(\beta)$$

where R is separation distance of the point collision partners, α is the polarizability of the neutral partner, μ is the dipole moment of the neutral partner, β is the angle between the dipole moment vector and the separation vector, and q is the charge on the ion.

All calculations are performed in the spherical coordinate system shown in Figure 1. The reactions are assumed to be a light particle (a deuterium atom) transfer to a heavy particle, so a "modified spectator stripping" model[4] has been chosen to describe the reaction. In this model, the reactions are assumed to occur when the reactants close to within some arbitrary separation distance. The transfer of mass is then assumed to take place with no change in the individual momenta of the collision partners.

The parameters required to characterize the interaction potential (α, μ, and q, but excluding β) as well as the maximum impact parameter, the collision energy, and the number of trajectories to be computed are input parameters read by the program. All other variables are determined by random number generator within the program. The initial velocity of the ion is assumed to be parallel to the z-axis.

In agreement with the "modified spectator stripping" model, a reaction is assumed to have occurred and the time-of-flight determined if the incident ion approaches within a specified distance of the target molecule and successfully retre⸱ s from the target neutral a distance of 15 Å. No information has been determined about the scattering angle,
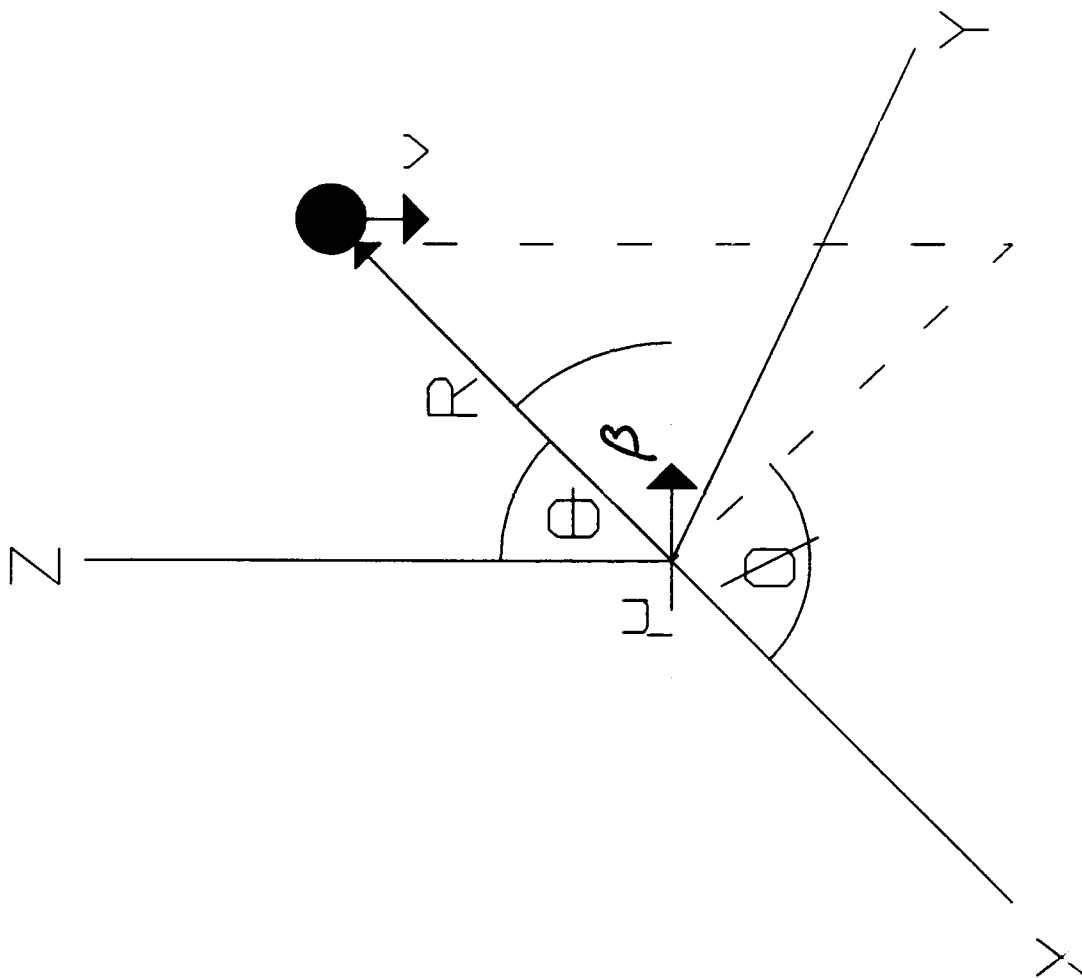
Figure 1. Coordinate system used in this study.

although the information may be obtained with a relatively simple modification to the program.

The program has been written in Microsoft FORTRAN (v.5.0) and run on a Zenith Data Systems Z-386/33 microcomputer equipped with an Intel 80387-33 math coprocessor.

III. RESULTS

The results of the calculations on the 29.5 Ev collision of $H_2{}^{18}O^+$ + D_2O are illustrated in Figures 2-5, and summarized in Table 1. A distinct decrease in both the location of the maximum of the time-of-flight distribution and the width of the time-of-flight distribution as the reaction distance decreases is observed. The best fit of the peak of the distribution is observed in the closest approach case studied (1 Å reaction distance). However, even in this case the time-of-flight distribution is significantly broader than that observed in the experiment.
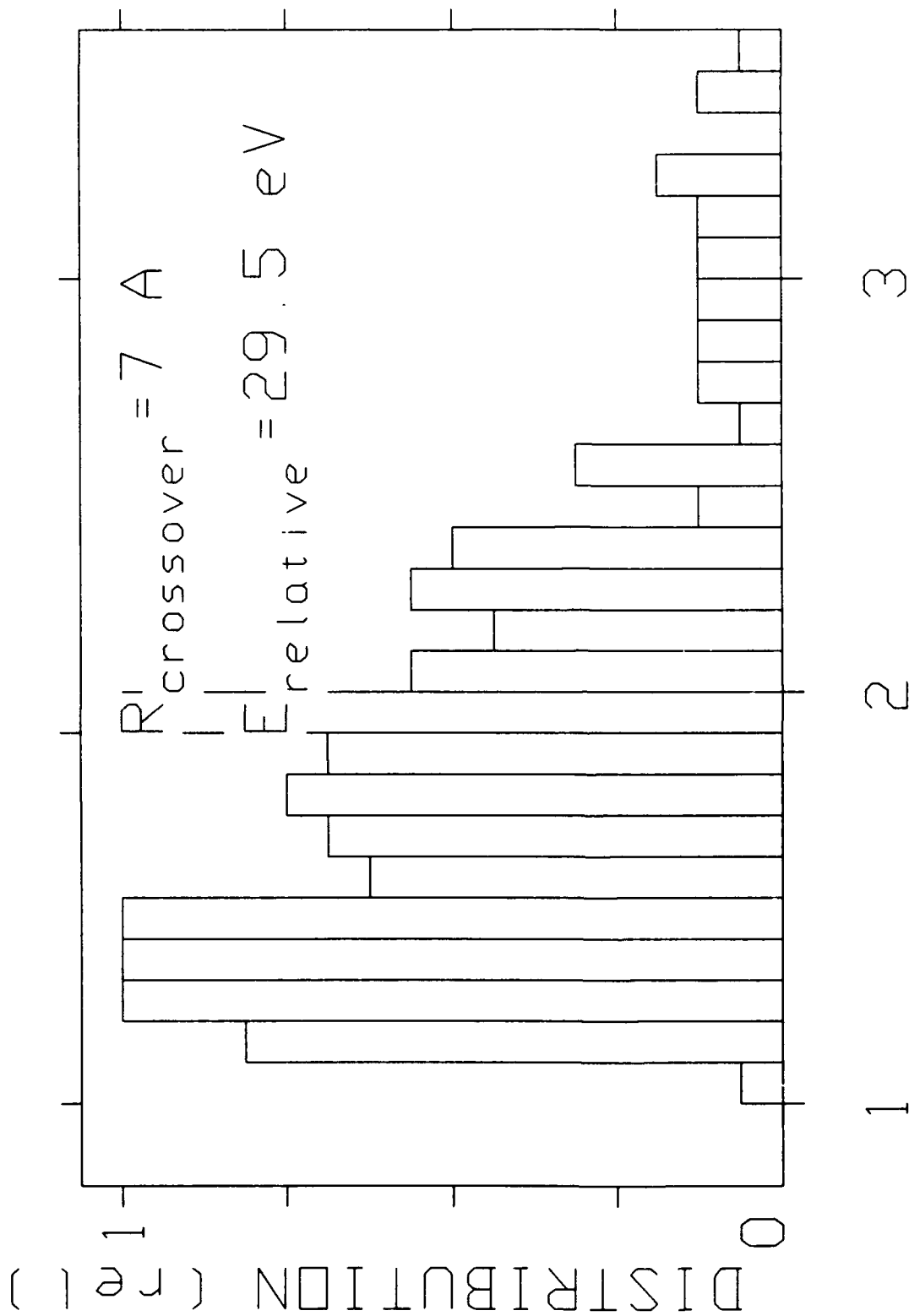
Figure 2. Time-of-flight Distribution relative to the fastest observed ion. $R_{crossover}$ is the reaction separation distance (7 Å).
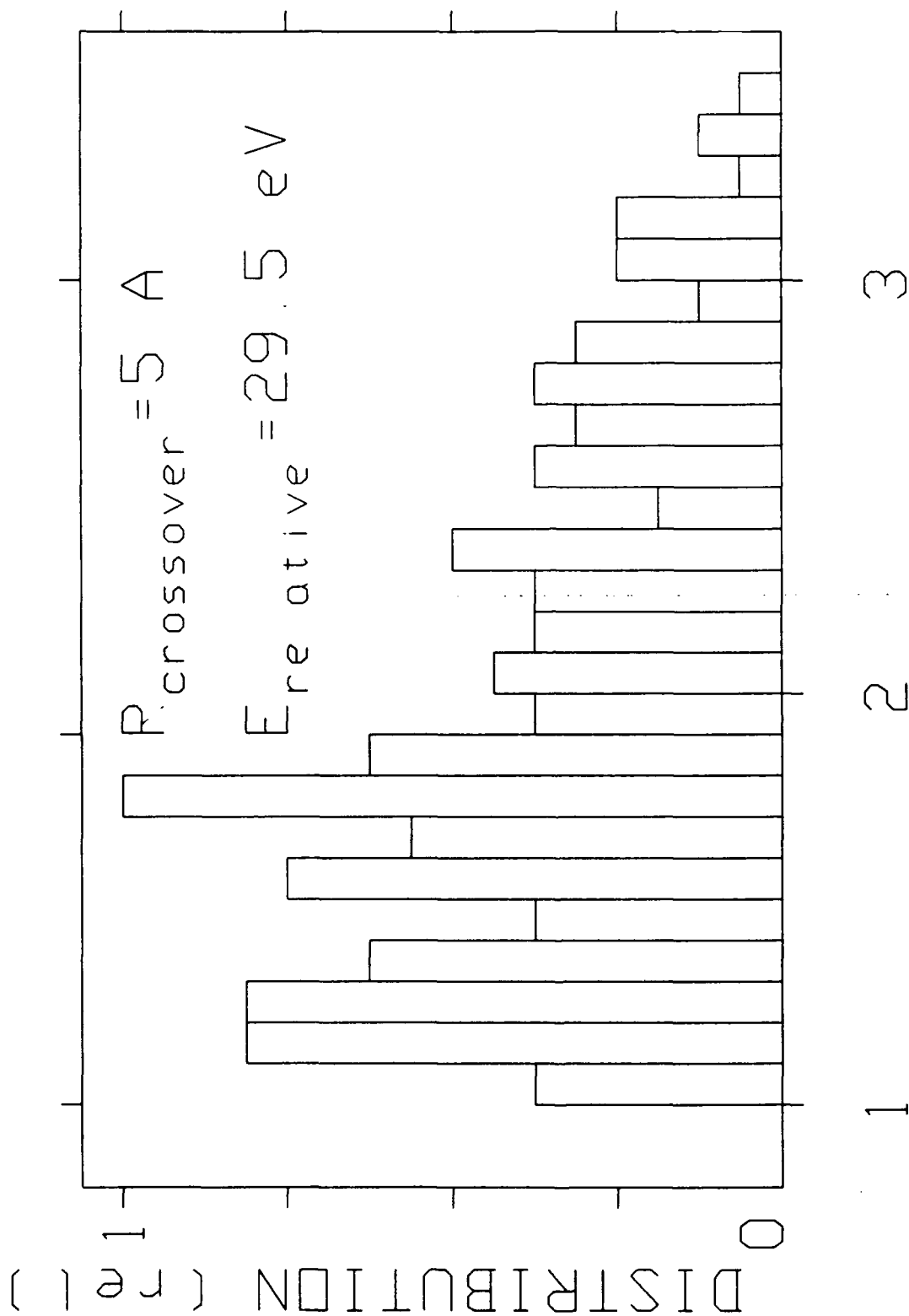
Figure 3. Time-of-flight Distribution relative to the fastest observed ion. $R_{crossover}$ is the reaction separation distance (5 Å).
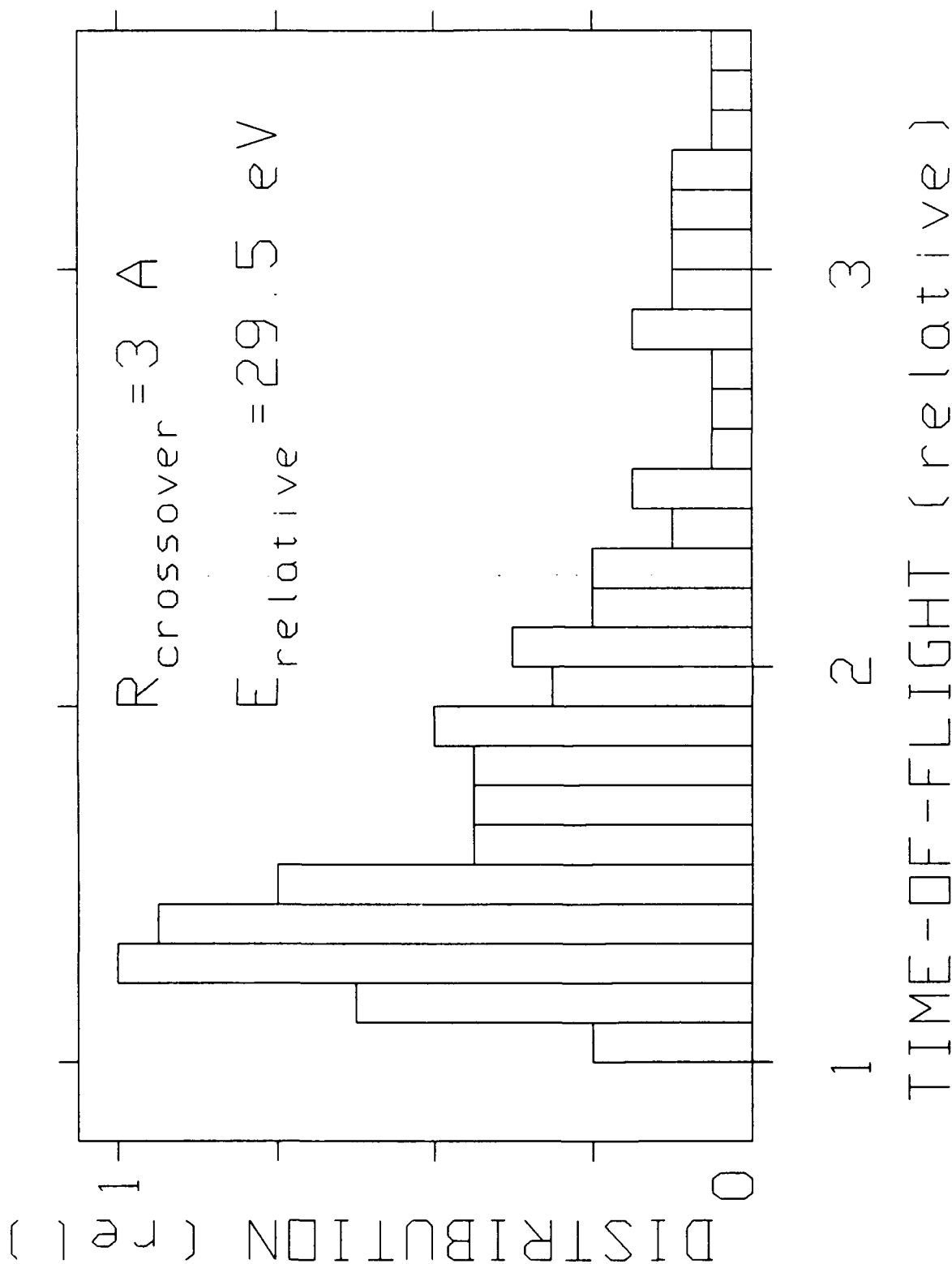
Figure 4. Time-of-flight Distribution relative to the fastest observed ion. $R_{crossover}$ is the reaction separation distance (3 Å).
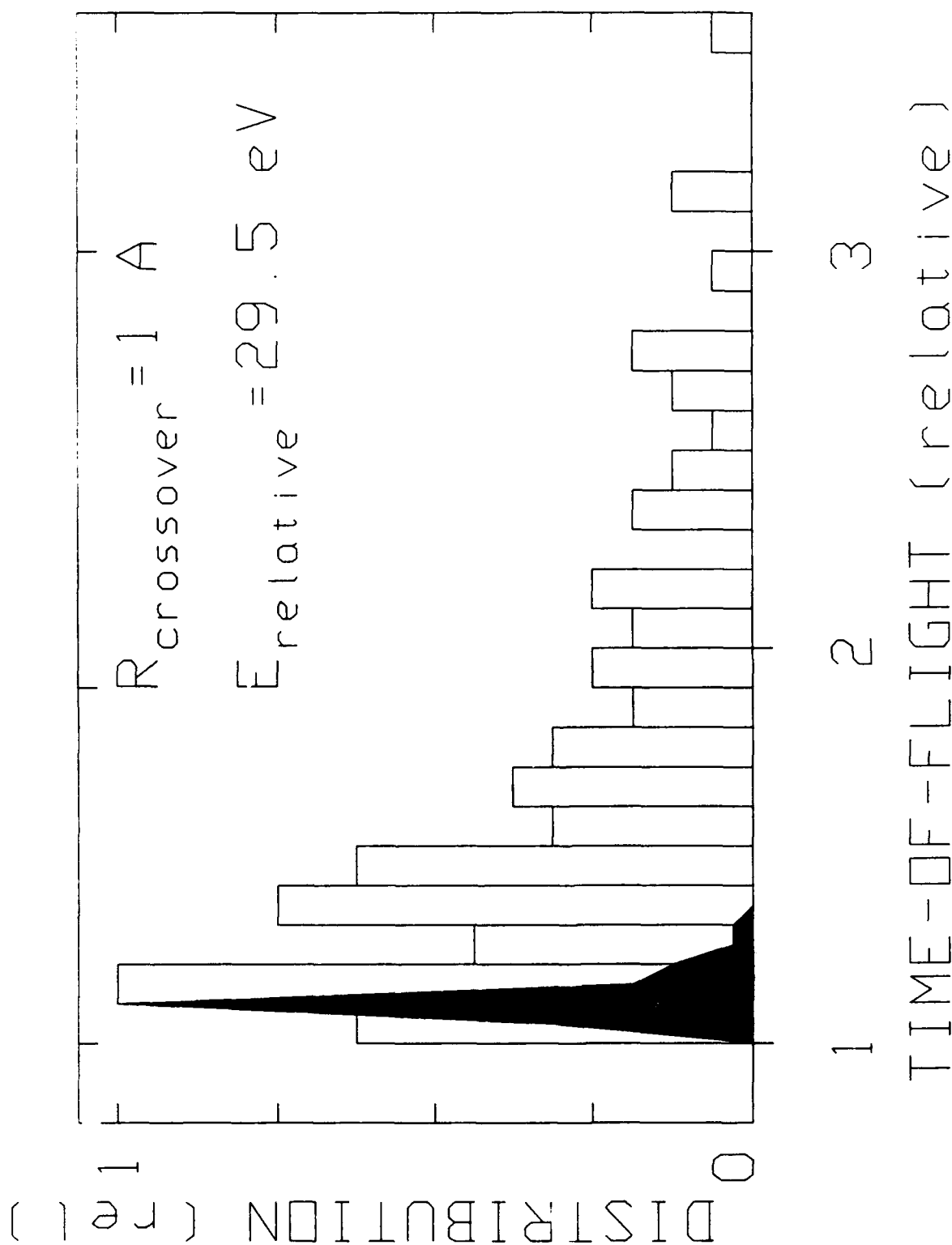
Figure 5. Time-of-flight Distribution relative to the fastest observed ion. $R_{crossover}$ is the reaction separation distance (1 Å). Shaded area is the experimental distribution.

Reaction Distance	Distribution Maximum	Distribution Width	Cross Section
(Å)	(rel)	(rel)	(Å^2)
7.0	1.4	1.3	20.0
5.0	1.3	0.8	16.8
3.0	1.2	0.7	14.3
1.0	1.1	0.5	3.95
Experiment	1.1	0.1	0.15

Table 1. Summary of results for $H_2^{18}O^+ + D_2O \rightarrow DH_2^{18}O^+ + OD$.

The reaction cross section can be calculated from expression

$$\sigma = \frac{N_{reactive}}{N_{total}} \pi \, (b_{max})2$$

where $N_{reactive}$ is the number of reactive trajectories observed, N_{total} is the total number of trajectories calculated, and b_{max} is the maximum permissible impact parameter.

IV. CONCLUSION

From the results reported above, it appears that the "modified spectator stripping" model can provide an adequate description of suprathermal energy ion-molecule collisions involving the transfer of light particles. Further work on this topic will include a study of the effect of limiting the observable scattering angle to only forward scattered events. This is expected to significantly reduce both the width of the time-of-flight distribution and the simulated cross section. Other reaction channels will also be examined to extend the results reported above.

V. REFERENCES

1. C.R. Lishawa, R.A. Dressler, J.A. Gardner, R.H. Salter, and E. Murad, <u>Cross sections and product kinetic energy analysis of H_2O^+-H_2O collisions at suprathermal energies.</u>, J. Chem. Phys. <u>93</u>(5), 1 September 1990, pp. 3196-3206.

2. R.A. Dressler, J.A. Gardner, C.R. Lishawa, R.H. Salter, and E. Murad, <u>Chemiluminescence Measurements of the N_2^+,N^+ + H_2O charge transfer systems at suprathermal energies</u>. J. Chem. Phys. <u>93</u>(12), pp. 9189-9191(L).

3. B. Carnahan, H.A. Luther, and J.O. Wilkes, "Applied Numerical Methods", John Wiley & Sons, New York, 1969.

4. Z. Herman, J. Kerstetter, T. Rose, and R. Wolfgang, <u>Disc. Faraday Soc.</u>, <u>44</u>, 123, (1967).

Electron Attachment to Transition-Metal Acids

Final Report

11 January - 31 December 1990

UES Project 210

Purchase Order No. S-210-10MG-113

Contract No. F49620-88-C-0053/SB5881-0378

University of Oklahoma Project 158-235

Submitted by: Thomas M. Miller

Department of Physics and Astronomy

University of Oklahoma

Norman, OK 73019

(617) 377-4028

Submitted to: Susan K. Espy

Contract Administrator

Universal Energy Systems, Inc.

4401 Dayton-Xenia Rd.

Dayton, OH 45432-1894

(513) 426-6900

Date: 21 December 1990

ABSTRACT

Experimental research on negative ions is described in this report. Two projects begun under the UES Summer Faculty Research Program were completed: (a) reactions of $Fe(CO)_n^-$ ions with methyl halides molecules were studied in the gas phase, for $n = 0$-4; (b) the gas-phase acidity of FeH was determined from proton transfer observations. A third, unforeseen, project was undertaken to aid in modeling a U. S. Air Force chemical release in the ionosphere; the electron affinity and electron attachment rate for SF_4 were measured, and reaction rates for SF_4^- interacting with various molecules were measured. Finally, electron attachment rates have been measured for several of the transition-metal hydrides for which we had earlier determined gas-phase acidities, under the SFRP.

1. Introduction and Objectives

The current Research Initiation Program (RIP) effort had two distinct aims. The first was to finish experimental work begun under the UES Summer Faculty Research Program (SFRP) during the summer of 1989, and to write up for publication the results of this work. The second goal was to install a new mass spectrometer on a flowing-afterglow Langmuir-probe (FALP) in order to determine the ion products of dissociative electron attachment to various transition-metal compounds, including transition-metal hydrides being synthesized by Amy E. S. Miller of the University of Oklahoma. Electron attachment rate coefficients are relatively straightforward to measure with the FALP apparatus.

2. Completion of SFRP Work

Further running time on the Geophysics Laboratory ion-molecule flow reactor could not be scheduled until the summer of 1990, but all data are now in hand completing a study of the gas-phase reaction rates and product branching ratios for iron and iron-carbonyl anions interacting with the methyl halides. This work is interesting for several reasons. It is the first study of a successively ligated transition-metal anion reacting with any molecule, $i.e.$, $Fe(CO)_n^-$, n = 0-4, were reacted. In particular, it is the first time, to our knowledge, that a reaction rate coefficient has been measured for the bare iron anion interacting with any molecule. Different product channels were seen: nucleophilic displacement, ligand exchange, halogen atom abstraction, and association. An estimate of the iron-methyl bond energy (1.1 eV or 25 kcal/mol) was obtained for $Fe(CO)CH_3$. Improvements of a draft of a paper on this research have been written under the present UES-RIP project, and the finished paper will be submitted to the *Journal of the American Chemical Society* within the coming months.

A second project, a study of reactions of Fe^- with various acids in the gas phase, has also been completed during the summer of 1990 except for the written report. The goal of this research was to bracket the gas-phase acidity of FeH by determining proton transfer rates between Fe^- and reference acids. A by-product of this work is additional data on reaction rates and product branching ratios for interactions of bare iron anions.

3. Electron Attachment

The primary apparatus improvement needed for a meaningful study of electron attachment to transition-metal compounds - and, indeed, to most molecules - was a good mass spectrometer. An Extrel quadrupole mass spectrometer is now in routine use to analyze the ionic products of the attachment. Equally important, but more mundane, has been the addition of high-quality MKS gas flow controllers and Baratron pressure sensors to the apparatus. Numerous tests have been performed using standards of both rapid and slow electron attachment.

The FALP apparatus was moved from the University of Oklahoma to the Geophysics Laboratory on 1 June 1990 so that full time could be devoted to this project, and so that studies which required close collaboration between the FALP and the GL selected-ion flow-tube (SIFT) could more easily be carried out.

Our work with the transition-metal hydrides is one example of such a collaboration. The GL-SIFT apparatus was used in the summer of 1989, under the UES-SFRP, to bracket the gas-phase acidities of several transition-metal hydrides, in relation to other extremely strong acids. The cobalt hydride, $HCo(PF_3)_4$, proved to be the strongest gas-phase acid known [see A. E. S. Miller, A. R. Kawamura, and T. M. Miller, *J. Amer. Chem. Soc. 112*, 457 (1990)]. The U. S. Air Force has applications for strong acids if they prove to attach electrons rapidly, yielding a very stable negative ion in the process. Thus, our FALP-based electron attachment results for $HCo(PF_3)_4$ take on added importance as a result of the SIFT-based research.

We have obtained electron attachment data for $HCo(PF_3)_4$ which indicate an attachment rate coefficient of 2 x 10^{-8} cm^3/s (*i.e.*, attachment occurs on about 1 in every 10 collisions). The mass spectra show that electron attachment to $HCo(PF_3)_4$ yields $Co(PF_3)_4^-$ in about 30% of the reactions and $HCo(PF_3)_3^-$ in the remaining reactions. The rate coefficient stated here is tentative; the mass spectra also show a few percent of $HCo(CO)(PF_3)_3$ impurity in the sample we studied. Further work, with a pure sample, is planned.

Similarly, we have obtained data for $HMn(CO)_5$, and find an attachment rate of 1.85 x 10^{-7} cm^3/s, with only the ion product $HMn(CO)_4^-$ observed. These electron attachment data are shown below. Other transition-metal hydride samples have been

prepared, and await time on the FALP. An abstract, titled "Gas-Phase Electron Attachment to Transition-Metal Hydrides", on this work has been submitted for a presentation for the Atlanta meeting of the American Chemical Society, 15-19 April 1991.

HMn(CO)$_5$
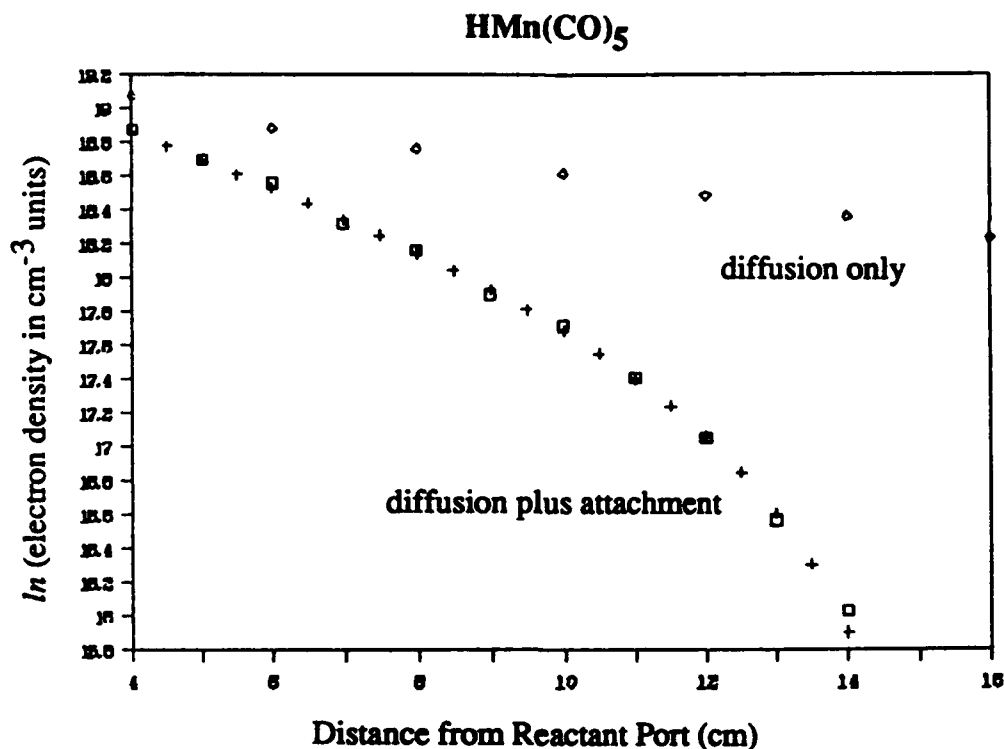
For HRe(CO)$_5$ we find an electron attachment rate of 1.32×10^{-7} cm^3/s.

We are seeking to understand periodic-table trends and to define the effect of ligand type upon gas-phase acidity and electron attachment efficiency. In the process, the relationship between acidity and attachment efficiency should emerge. An interpretation of the SIFT/FALP work with the transition-metal hydrides is obviously premature at this point. SIFT proton transfer data are needed for a few of the hydrides that had not been synthesized in time for our work in the summer of 1989. Further FALP data are needed for all these compounds.

Other examples of the SIFT/FALP collaboration have arisen, to no surprise. Chemical modeling of U. S. Air Force releases of SF$_6$ vapor in the earth's ionosphere requires data not only for the primary SF$_6$ interaction with the ionospheric plasma, but also for all products of these and subsequent reactions. During December 1990 we have focused on SF$_4$, for the following reason. The electron affinity of SF$_4$ was presumed

known (at 2.35 eV) from charge-transfer bracketing reactions carried out with reference molecules, at Los Alamos National Laboratory (LANL) in the early 1980's. In light of today's very accurate electron affinities for the reference molecules used in that study, however, we realized that the brackets were contradictory, and at least one of the reported reactions was incorrect. As we repeated these earlier measurements, using the SIFT, we found that most of the published results for anions reacting with SF_4 were incorrect. We have now determined rate coefficients and product branching ratios for 31 different reactions involving SF_4 or SF_4^-. Many of these reaction results will be used as input data for the modeling codes. On the basis of this work, we now believe that the electron affinity of SF_4 is near 1.65 eV.

Concurrently, the FALP apparatus was used to study electron attachment to SF_4. An attachment rate of 1.57×10^{-9} cm^3/s was measured. (This rate implies that attachment occurs only on about one of every hundred encounters between an electron and an SF_4 molecule.) This measured rate agrees with that determined earlier at LANL. However, the mass spectrum from the FALP shows that considerable SF_6^- (0.46%) is present in addition to the expected SF_4^-. A quick calculation showed that it would only take an 0.7% SF_6 impurity in the SF_4 gas to completely account for the attachment observed, since SF_6 attaches rapidly - and the gas supplier can only say that the SF_6 impurity is "less than 3%". It was clear from the SF_4^- observed that SF_4 indeed attached electrons, but the task that remained was to determine exactly what fraction of the FALP plasma decay was due to SF_4 alone. We performed the following experiment: the initial electron density was made large enough to convert the SF_6 impurity rapidly into SF_6^- within a few centimeters of the reactant port. The FALP plasma decay was then solely dependent upon SF_4, beyond the point of total SF_6 loss. An electron attachment rate coefficient of 1.4×10^{-8} cm^3/s was measured for SF_4, which is a factor of ten greater than the present literature value. Additional experiments are planned to confirm this result.

4. Recommendations

The temperature variation of the electron attachment rate and of the product branching ratios are of scientific interest because the variation gives an indication of how the overlap of neutral and anion vibrational wavefunctions affects attachment. High temperature data are of great practical importance to the U. S. Air Force. Thus, we have designed a variable temperature flow tube for the FALP. Materials are on order; the new

flow tube will be constructed at the University of Oklahoma machine shop and shipped to GL in the early spring of 1991.

5. Acknowledgments

1989-1990 USAF-UES RESEARCH INITIATION PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by

UNIVERSAL ENERGY SYSTEMS, INC.

FINAL REPORT

CO2(4.3μm) VIBRATIONAL TEMPERATURES AND LIMB RADIANCES

IN THE MESOSPHERE AND LOWER THERMOSPHERE:

SUNLIT CONDITIONS AND TERMINATOR CONDITIONS

Principal Investigator:	Dr. Henry Nebel
	Professor of Physics
	Alfred University
	Alfred, NY 14802
Research Location:	Geophysics Laboratory/Optical & Infrared Technology Division/ Infrared Backgrounds Branch Hanscom AFB, MA 01731
Date:	3 December 1990
Contract No:	F49620-88-C-0053/SB5881-0378

CO2(4.3μm) VIBRATIONAL TEMPERATURES AND LIMB RADIANCES

IN THE MESOSPHERE AND LOWER THERMOSPHERE:

SUNLIT CONDITIONS AND TERMINATOR CONDITIONS

by

Henry Nebel

ABSTRACT

Vibrational temperature profiles as functions of altitude under both sunlit and terminator conditions have been calculated for the 4.3 μm, 2.7 μm and 2.0 μm absorbing states of carbon dioxide which contribute to 4.3 μm radiation. These calculations have been performed using a non-equilibrium line-by-line infrared radiation transport code (ARC) developed at the Geophysics Laboratory. The vibrational temperature profiles are then used to calculate integrated radiance from the 4.3 μm bands of CO_2 in a limb view for the mesosphere and lower thermosphere. The calculated radiances are in good agreement with measurements obtained by the Spectral Infrared Rocket Experiment (SPIRE)* for the sunlit case, but underestimate the measured radiances in the terminator case.

*Stair, A.T., et. al., J. Geophys. Res. 90, 9763-9775 (1985).

ACKNOWLEDGEMENTS

INTRODUCTION:

There has been increasing interest recently in radiative transfer in the infrared spectral region under non-equilibrium conditions in the atmosphere (1-5). Under these conditions, local thermodynamic equilibrium (LTE) may not be assumed to apply, i.e. collisions among molecules are not frequent enough to bring a parcel of air into equilibrium before radiative deexcitation occurs. This is generally the situation in the earth's upper atmosphere (above 60 km). An infrared radiance computer code has been developed at the Geophysics Laboratory (GL) by Dr. Ramesh Sharma of GL and Dr. Peter Wintersteiner of Arcon Corporation. This code treats absorption, emission, and transmission of infrared radiation through the atmosphere under non-equilibrium conditions. One component of the code (RAD) calculates excited state population densities and the corresponding vibrational temperature profiles assuming various mechanisms of excitation and deexcitation, and also assuming the population densities are constant in time. Another component of the code (NLTE) uses the vibrational temperature profiles to calculate total integrated band radiance for various viewing geometries. These calculations are done on a line-by-line basis, whereas most previous calculations of this type were based on band models (1,2,5).

OBJECTIVES:

The code has previously been applied to the 4.3 μm band of carbon dioxide under night time conditions in the upper atmosphere (6,7). The calculated vibrational temperature profiles are in good agreement with others based on band models (2,5). The code has also been applied to the same band of CO_2 under day time (sunlit) conditions (8). The resulting limb view integrated radiances are in excellent agreement with day time measurements obtained by the Spectral Infrared Rocket

Experiment, or SPIRE (9). The present report describes further work on this problem to accomplish the following objectives: (A) a re-calculation of the sunlit case using a coupled version of the RAD code and a more realistic CO_2 altitude profile; (B) an attempt to treat the 4.3 μm band of CO_2 under terminator conditions where the limb view line-of-sight is partly sunlit and partly in darkness due to the sun being below the horizon. The rationale for doing this is the fact that the SPIRE mission took a number of terminator scans. If these can be modelled, then results of theoretical calculations may be compared with existing experimental data.

APPROACH:

Objective (A): An energy level diagram including the vibrational states of CO_2 required for this calculation is shown in Fig. 1. The notation for the vibrational levels is that of the HITRAN database generated at the Geophysics Laboratory (10). The first, second and fourth digits represent the three quantum numbers of the vibrational modes of excitation of CO_2 as a linear triatomic molecule. The 4.3 μm absorption or emission band results from transitions between states with one ν_3 quantum and corresponding states with no ν_3 quanta (other quantum numbers being equal for the upper and lower states). Vibrational temperature profiles for the 00011 state have been obtained for the principal isotopic form of carbon dioxide ($^{12}C^{16}O_2$) labeled 626, as well as for the three most important minor isotopic forms ($^{13}C^{16}O_2$, $^{16}O^{12}C^{18}O$, $^{16}O^{12}C^{17}O$) labeled 636, 628, and 627 respectively. Previously, the vibrational temperatures were calculated for each isotopic form separately. The RAD program has now been re-written in such a way that the 00011 states for all 4 of these isotopic forms are coupled together. Thus the vibrational temperatures for all 4 states are determined simultaneously along with that of the first excited N_2 state, which is closely coupled with the others because it has about the same energy. This has been done for night

time conditions and also for sunlit conditions assuming a solar elevation angle of 12°
which is representative of the sunlit scans of the Spectral Infrared Rocket Experiment
(9). Vibrational temperatures for the 01111 state are calculated using program RAD for
the 626 and 636 isotopes separately, and for the higher-lying states in Fig. 1 using a
simpler program called VPMP which neglects the airglow contribution compared to
solar absorption.

For each state, vibrational temperatures are used as input to program NLTE
which calculates total band radiance in a limb view, the geometry of which is shown in
Fig. 2. Results for all states considered are added and then compared with radiances
measured by the SPIRE mission as a function of tangent height (see Fig. 2).

The experimental 4.3 μm radiance values as measured by SPIRE, integrated
between 4.12 and 4.49 μm, are shown in Fig. 3. Scans 9-12 are the sunlit scans referred
to above, while scans 1-8 are scans which were made across the dawn terminator,
where the limb view line-of-sight was partly sunlit and partly in darkness. Data from
these scans is referred to as "night-time" in reference 9. The great disparity between
these two data sets emphasizes the importance of the role that solar pumping plays in
exciting the CO_2 states that emit at 4.3 μm.

Objective (B): Solar flux absorption coefficients could previously be calculated
only for solar zenith angles less than 90° (sun above the horizon), resulting in too
much solar absorption for a below-horizon sun. This limitation has now been
overcome. Solar flux absorption coefficients have been calculated for a solar zenith
angle of 95° (sun 5° below horizon) which is representative of the terminator ("night-
time") scans of the SPIRE mission. Vibrational temperature profiles have been
calculated for this below-horizon sun using the coupled form of program RAD
described above. A modified version (NLTEA) of program NLTE is then used which is
capable of dividing the viewing path into two segments characterized by two different
sets of vibrational temperatures. In other words, the vibrational temperatures for the

dark and sunlit portions of the line-of-sight are modelled separately. Night time vibrational temperatures (no solar absorption) are used for the dark part of the path, and vibrational temperatures for a solar zenith angle of 95° are used for the sunlit part of the path. Similar calculations are performed for some of the higher-lying excited states shown in Fig. 1, and the results are added to the radiances calculated for the fundamental (00011 state). The final radiances may then be compared with terminator measurements obtained by the Spectral Infrared Rocket Experiment (see Fig. 3).

RESULTS:

(A) Vibrational temperature profiles for the 00011 state for all four isotopic forms of CO_2 are shown in Fig. 4 for the sunlit case described previously. These profiles are calculated using the coupled form of program RAD referred to above, and a CO_2 profile based on measurements obtained from rocket-borne instruments. It is believed that this profile is more reliable than profiles used previously. Vibrational temperature profiles for the 01111 state for the 626 and 636 isotopic forms are shown in Fig. 5. Similar profiles (not shown) have also been calculated for the higher-lying states shown in Fig. 1 using program VPMP. Total band radiance in a limb view from all states considered is shown in Fig. 6 as a function of tangent height (see Fig. 2), together with measured values obtained under sunlit conditions by the Spectral Infrared Rocket Experiment (9). As can be seen in Fig. 6, the agreement is excellent for 50 to 100 km tangent heights, but the calculated values are somewhat lower than the measurements above 105 km. The discrepancy is of the same order of magnitude as the nominal noise of the SPIRE instrument, which would be about 4×10^{-9} watt/cm^2-sr for the 4.3 μm band, the effect of which can be seen in the scatter of the terminator data above 105 km (see Fig. 3). This noise was not subtracted from the integrated spectral data shown in Fig. 3. If this noise were subtracted, there would be better agreement between model

predictions and data. In addition, the total band radiance at these high altitudes depends almost directly on the CO_2 concentration assumed in the atmospheric profile used in the calculations. Since this quantity is not well known at the upper altitudes, a discrepancy between calculated and measured radiance values in this region is not surprising.

Fig. 7 shows the breakdown of the daytime model radiance into contributions from individual bands or groups of bands. Several features are noteworthy: (1) The major isotope ν_3 fundamental (00011 626) plays a relatively unimportant role for tangent heights in the mesosphere (50-90 km). (2) The bands originating in the states lying near 3500 cm^{-1} (Group 1 states in Fig. 1) provide the major contribution to the total radiance for tangent heights of 55-95 km. These states absorb sunlight at 2.7 μm and then emit radiation preferentially at 4.3 μm. (3) For the lower tangent heights, the contributions from the various groups of bands considered are roughly comparable.

(B) Night time vibrational temperature profiles for the 00011 state as calculated by the coupled version of program RAD are shown in Fig. 8. Similar curves including solar absorption at a solar zenith angle of 95° are shown in Fig. 9. Vibrational temperature profiles (not shown) for some of the higher-lying states of Fig. 1 have also been calculated for this solar zenith angle using program VPMP. Total band radiance in a limb view has been calculated for terminator conditions by using the night time vibrational temperatures (Fig. 8) for the dark portion of the line-of-sight and the 95° vibrational temperatures (Fig. 9) for the sunlit portion. Results are shown in Fig. 10 as a function of tangent height, along with the terminator ("night-time") measurements obtained by the Spectral Infrared Rocket Experiment. As can be seen in this figure, the calculations under-estimate the measured radiances by a substantial factor for tangent heights above 60 km. One possible explanation for this discrepancy is the following: For low sun conditions, and particularly when the solar zenith angle is greater than 95°, lateral variations in the solar flux absorption caused by variations of the solar

zenith angle along the line-of-sight are much greater than for high-sun cases. A one-dimensional model such as that used here cannot account for these variations, so the effect of net horizontal radiative transport (expected across the terminator) cannot be evaluated explicitly. The solution of this problem would be to calculate different vibrational temperature profiles for different elements along a given line-of-sight, in order to reflect the varying solar zenith angle along the path. This would be a rather laborious task which has not been attempted as yet.

Two additional mechanisms which might affect the calculated band radiance in the terminator case have been considered. One involves the hydroxyl radical OH. It is well known (11) that this radical exists in the upper atmosphere, and it has been suggested (12) that OH might excite nitrogen molecules which in turn might excite $CO_2(00011)$ according to the following reactions:

$$N_2(0) + OH^*(v < 9) \longrightarrow N_2(1) + OH^*(v-1) + E_v$$

$$N_2(1) + CO_2(0) \longrightarrow N_2(0) + CO_2(00011) + \Delta E.$$
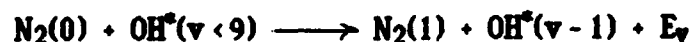
Inclusion of this mechanism in the calculations results in slightly increased band radiance in the 60-100 km altitude range, but not nearly enough to account for the experimental measurements.

The second mechanism considered involves electronically excited atomic oxygen $O(^1D)$. It is well known (13) that $O(^1D)$ is plentiful in the upper atmosphere, and it has been suggested (14) that $O(^1D)$ might also excite N_2 molecules leading to increased CO_2 concentrations according to the reactions

$$N_2(0) + O(^1D) \longrightarrow N_2(1) + O(^3P) + \Delta E$$

$$N_2(1) + CO_2(0) \longrightarrow N_2(0) + CO_2(00011) + \Delta E.$$
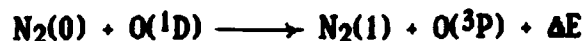
Inclusion of this mechanism in the calculations also results in slightly increased band radiance in the 60-100 km altitude range, but not nearly enough to account for the experimental results. The discrepancy between theory and experiment for this case has not yet been resolved.

CONCLUSIONS:

The infrared radiance computer code developed at the Geophysics Laboratory is an accurate and efficient vehicle for modelling infrared radiative transfer through the atmosphere under non-equilibrium conditions. Using the code to calculate vibrational temperature profiles and limb view integrated radiances for the 4.3 micron band of carbon dioxide under sunlit conditions results in excellent agreement with measurements obtained by the Spectral Infrared Rocket Experiment, or SPIRE (9), for 50-100 km tangent heights.

Applying the code to the terminator case where the sun is below the horizon results in calculated limb view integrated radiances which are significantly lower than measurements obtained by the SPIRE mission for 60-100 km tangent heights. One possible explanation of the discrepancy is the lateral variation in solar flux absorption along the line-of-sight with the possible effect of horizontal radiative transport across the terminator as discussed in the previous section. In any case, the discrepancy has not yet been resolved.

The US Air Force Research Initiation Program (sponsored by AFOSR, conducted by UES) has been very successful in allowing me to continue research begun while on the Summer Faculty Research Program during the summer of 1989. The work described in this report carries out all recommendations listed in my Final Report for the 1989 SFRP. These recommendations include using the coupled form of program RAD as discussed above, applying the code to the terminator case for comparison with SPIRE measurements, and considering possible effects of the hydroxyl radical (OH) in the calculations. Most of the work was performed during the summer of 1990 while in residence at the Geophysics Laboratory, Hanscom AFB, Mass. In addition, the RIP award provided funds for two hours of released time from teaching during each of the semesters Spring 1990 and Fall 1990. The continuity fostered by these research

programs has been very beneficial to me and to my research colleagues at the Geophysics Laboratory. I hope to continue this research under future sponsorship of AFOSR, working on problems of interest to the Air Force.

RECOMMENDATIONS:

The discrepancy between calculated and measured radiance values for the terminator case described above needs to be resolved. One possible refinement would be the calculation of $N_2(1)$ vibrational temperatures considering this state as being coupled to some of the higher-lying states as well as to the 00011 state of CO_2. This might result in additional excitation of the 00011 state and thus increased radiance. Another refinement would be the development of a two-dimensional model to describe the radiative transfer, and in this way account for possible horizontal radiative transport across the terminator as discussed earlier. However, this would represent an order-of-magnitude increase in the complexity of the code and the calculations. We hope to initiate such a project some time in the future.

It is hoped that CO_2 concentration profiles will be measured at high latitudes, in order to more accurately represent the conditions of the SPIRE experiment which was launched from Poker Flat, Alaska. We know of no such measurements which have been made at high latitudes. This, of course, would be a project for an experimental group, rather than a modelling group such as ours.

Finally, it is recommended that the Summer Faculty Research Program and the Research Initiation Program continue, since together they represent a very useful means of establishing and maintaining ongoing research relationships between college faculty and Air Force laboratory professionals.

REFERENCES

1. Kumer, J. B., and T. C. James, "$CO_2(001)$ and N_2 Vibrational Temperatures in the $50 \lesssim z \lesssim 130$ km Altitude Range", J. Geophys. Res. 79, 638-648 (1974).

2. Shved, G. M., G. I. Stepanova, and A. A. Kutepov, "Transfer of 4.3 μm CO_2 Radiation on Departure from Local Thermodynamic Equilibrium in the Atmosphere of the Earth", Izvestia, Atmospheric and Oceanic Physics, 14, 589-596 (1978).

3. Sharma, R. D., and P. P. Wintersteiner, "CO_2 Component of Daytime Earth Limb Emission at 2.7 Micrometers", J. Geophys. Res. 90, 9789-9803 (1985).

4. Solomon, S., J. T. Kiehl, B. J. Kerridge, E. E. Remsberg, and J. M. Russel III, "Evidence for Nonlocal Thermodynamic Equilibrium in the V_3 Mode of Mesospheric Ozone", J. Geophys. Res. 91, 9865-9876 (1986).

5. López-Puertas, M., R. Rodrigo, J. J. López-Moreno, and F. W. Taylor, "A non-LTE radiative transfer model for infrared bands in the middle atmosphere. II. $CO_2(2.7$ and 4.3 μm) and water vapor (6.3 μm) bands and $N_2(1)$ and $O_2(1)$ vibrational levels", J. Atmos. Terr. Phys. 48, 749-764 (1986).

6. Nebel, H., "Night-Time $CO_2(001)$ Vibrational Temperatures and Limb-View Integrated Radiances in the 50 to 150 km Altitude Range", Final Report/ 1987 USAF-UES Summer Faculty Research Program (1987).

7. Nebel, H., R. D. Sharma, and P. P. Wintersteiner, "Night-Time $CO_2(001)$ Vibrational Temperatures in the 50 to 150 km Altitude Range", Eos. Trans. AGU, 68, 1389 (1987).

8. Nebel, H., "CO$_2$(4.3µm) Vibrational Temperatures and Limb Radiances Under Sunlit Conditions in the 50-120 km Altitude Range", Final Report/ 1989 USAF-UES Summer Faculty Research Program (1989).

9. Stair, A. T., R. D. Sharma, R. M. Nadile, D. J. Baker, and W. F. Grieder, "Observations of Limb Radiance With Cryogenic Spectral Infrared Rocket Experiment", J. Geophys. Res. 90, 9763-9775 (1985).

10. Rothman, L. S., R. R. Gamache, A. Goldman, L. R. Brown, R. A. Toth, H. M. Pickett, R. L. Poynter, J.-M. Flaud, C. Camy-Peyret, A. Barbi, N. Husson, C. P. Rinsland, and M. A. H. Smith, "The HITRAN Database: 1986 Edition", Appl. Opt. 26, 4058-4097 (1987).

11. Baker, D. J., T. Conley, and A. T. Stair, "On the Altitude of the OH Airglow", Eos Trans. AGU, 58, 460 (1977).

12. Kumer, J. B., A. T. Stair, N. Wheeler, K. D. Baker, and D. J. Baker, "Evidence for an OH*$\xrightarrow{VV}$N$_2$*$\xrightarrow{VV}$CO$_2$(V_3')$\longrightarrow$CO$_2$ + hν(4.3 µm) Mechanism for 4.3 µm Airglow", J. Geophys. Res. 83, 4743-4747 (1978).

13. Slanger, T. G., and G. Black, "Electronic to vibrational energy transfer efficiency in the O(^{1}D)-N$_2$ and O(^{1}D)-CO systems", J. Chem. Phys., 60, 468-477 (1974).

14. Harris, R. D., and G. W. Adams, "Where Does The O(^{1}D) Energy Go?", J. Geophys. Res. 88, 4918-4928 (1983).
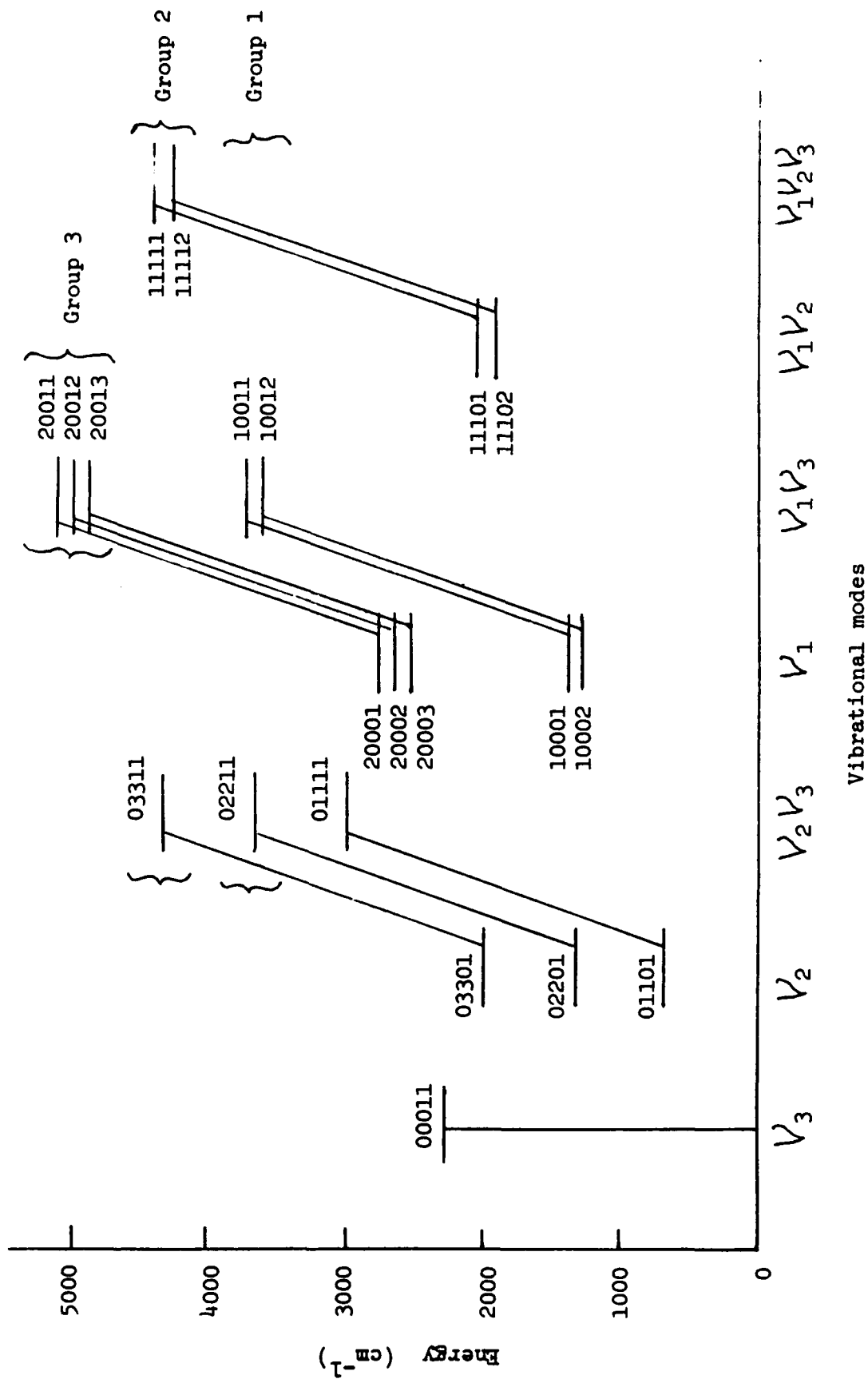
Fig. 1. Vibrational levels and 4.3 micron transitions considered.
Absorption at 2.7 and 2.0 microns not shown.
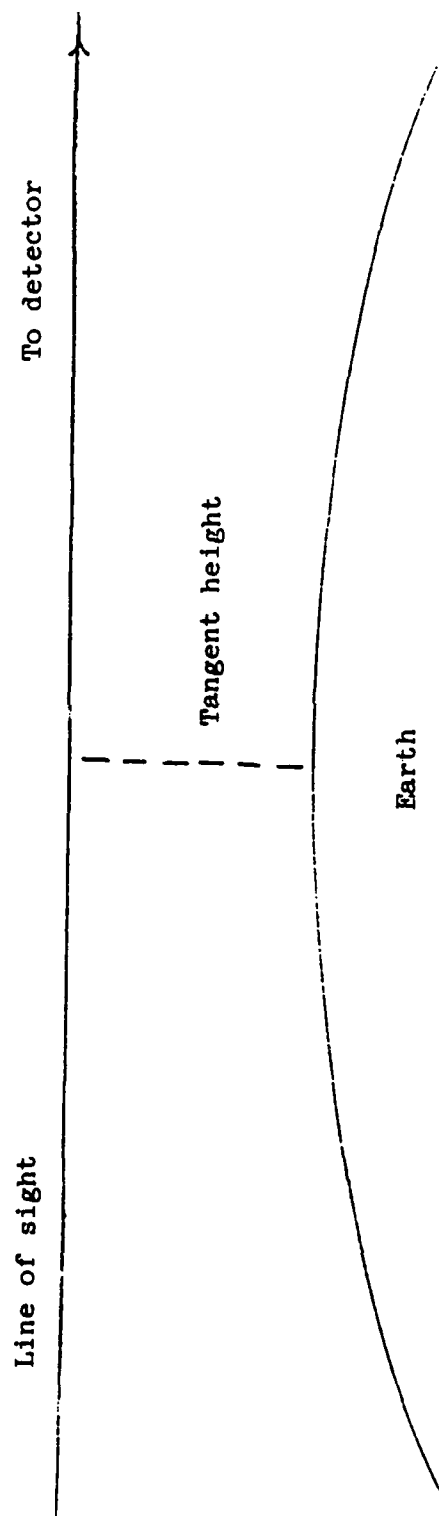
Fig. 2. Limb viewing geometry.
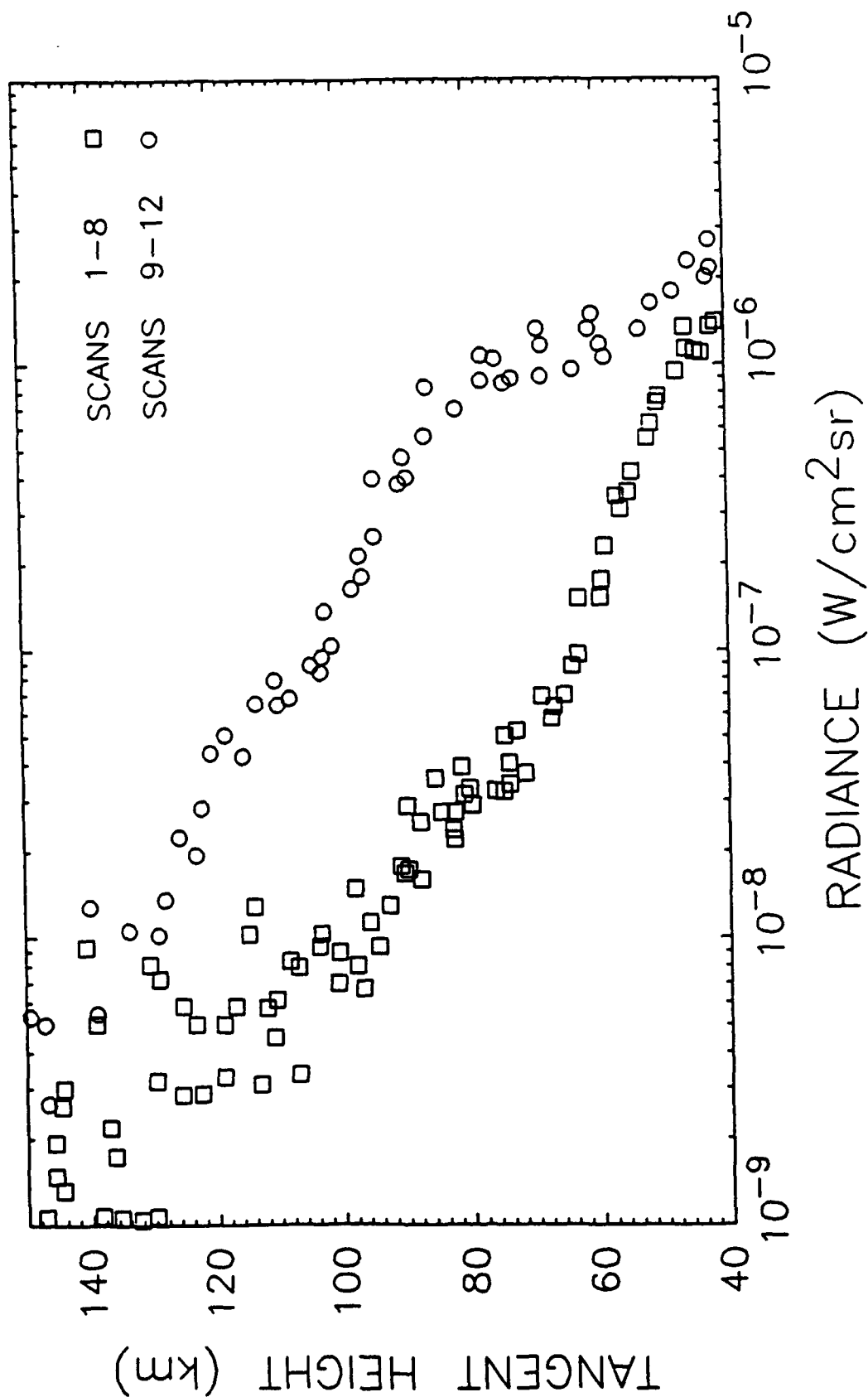
Fig. 3. Integrated 4.3 micron radiance from SPIRE.

00011 STATE DAYTIME

SZA = 78°

Temp

636

626

627

628

VIBRATIONAL TEMPERATURE (K)

240 260 280 300

ALTITUDE (km)

40 60 80 100 120 140
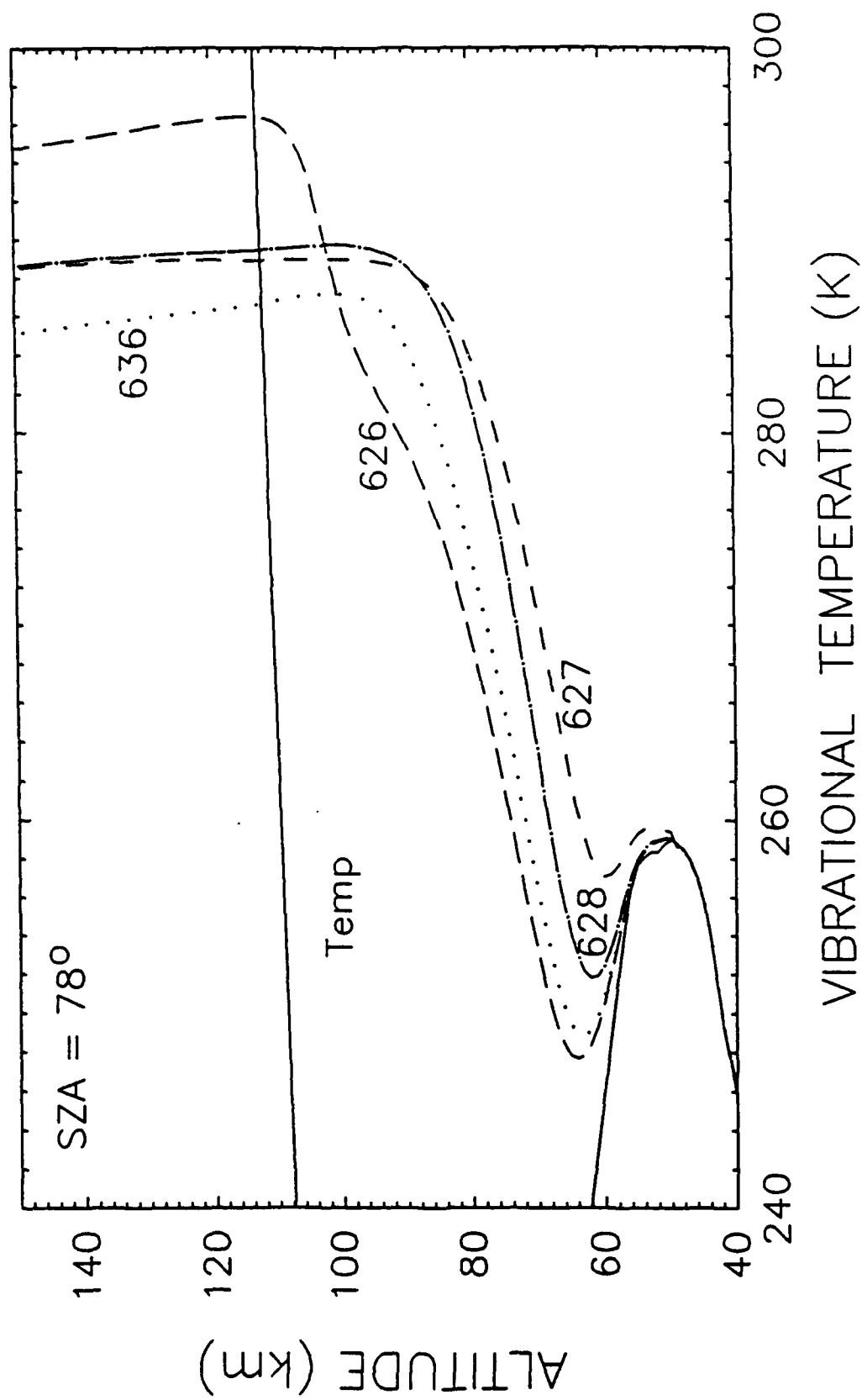
Fig. 4. Vibrational Temperatures for 00011 state.

01111 STATE DAYTIME

SZA = 78°

Temp

636

626

VIBRATIONAL TEMPERATURE (K)

ALTITUDE (km)
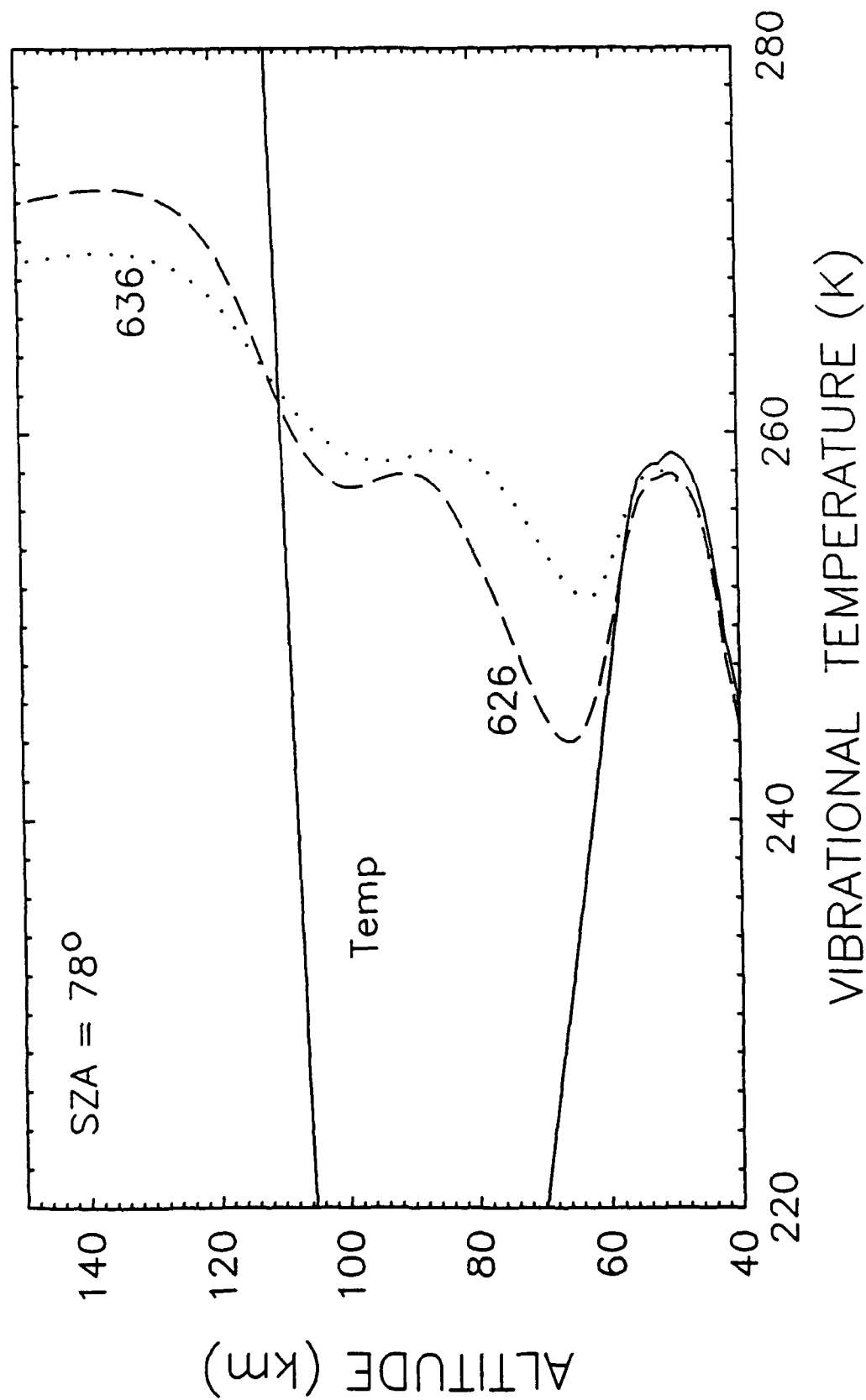
Fig. 5. Vibrational Temperatures for 01111 state.
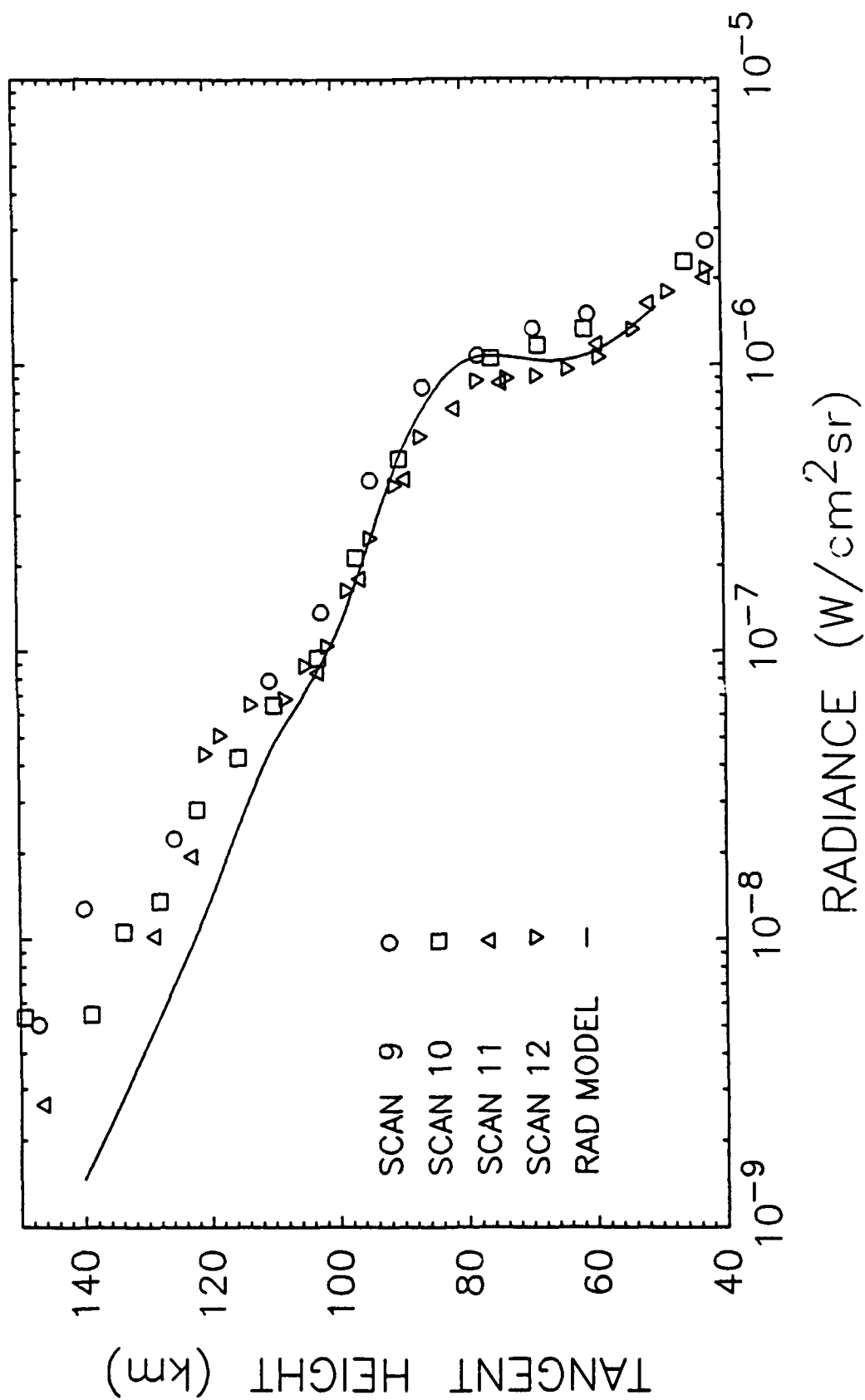
Fig. 6. Total 4.3 micron radiance in a limb view (sunlit).
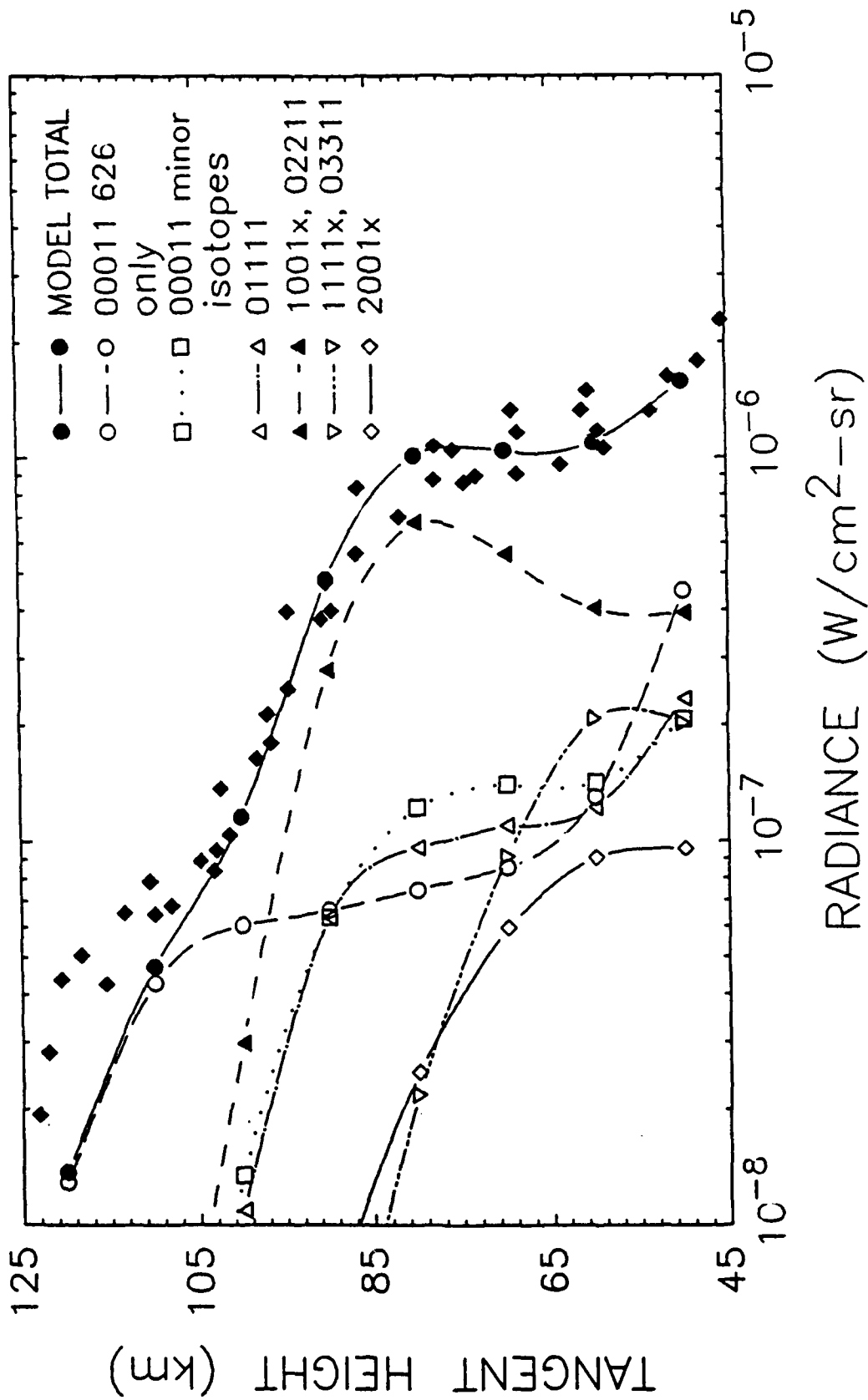
Fig. 7. Contributions to sunlit 4.3 micron radiance.

00011 STATE NIGHTTIME

Temp

627
628
636
626

VIBRATIONAL TEMPERATURE (K)

ALTITUDE (km)
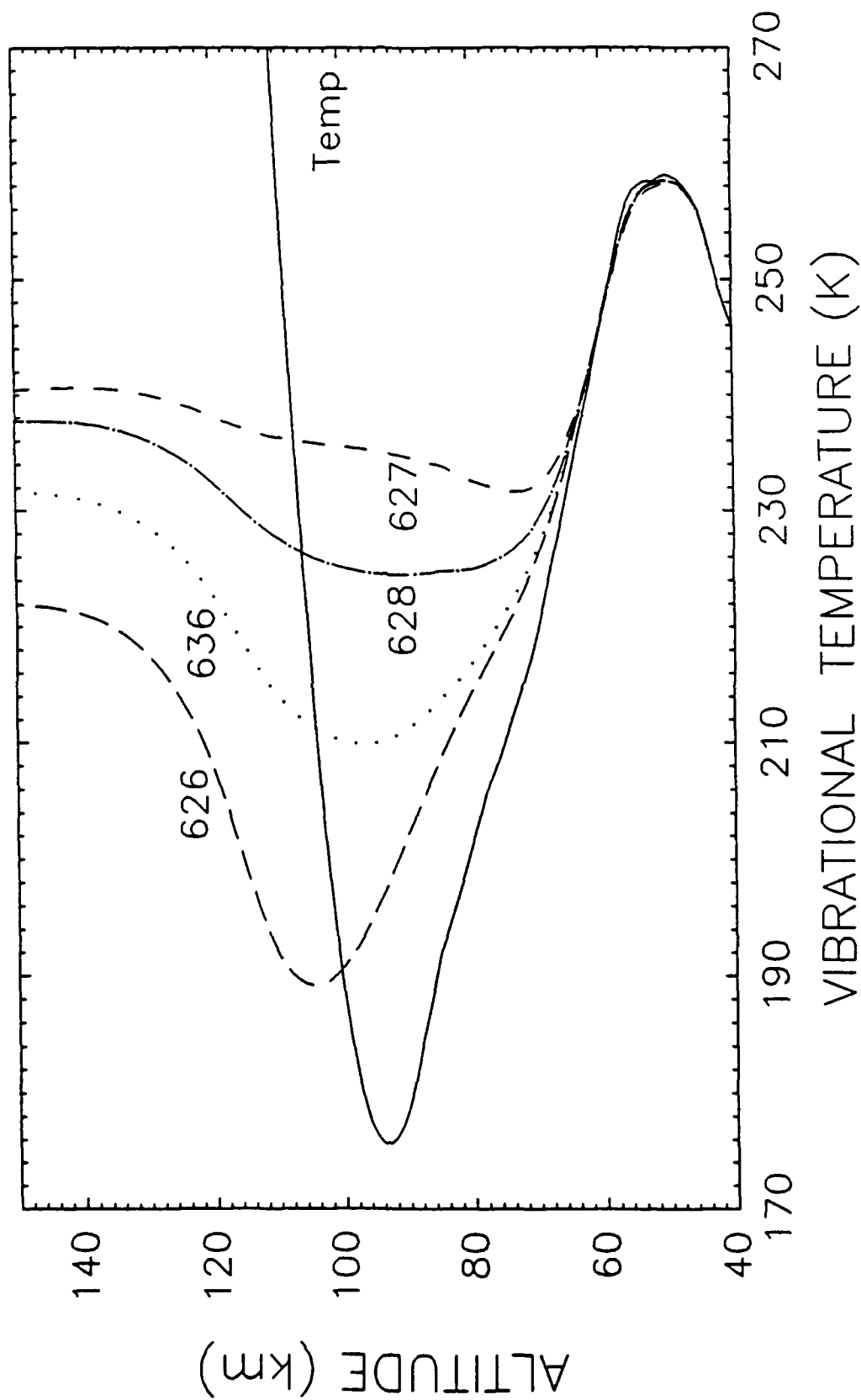
Fig. 8. Night-time Vibrational Temperatures for 00011 state.

00011 STATE DAYTIME

SZA = 95°

636

626

627

628

Temp

VIBRATIONAL TEMPERATURE (K)

ALTITUDE (km)
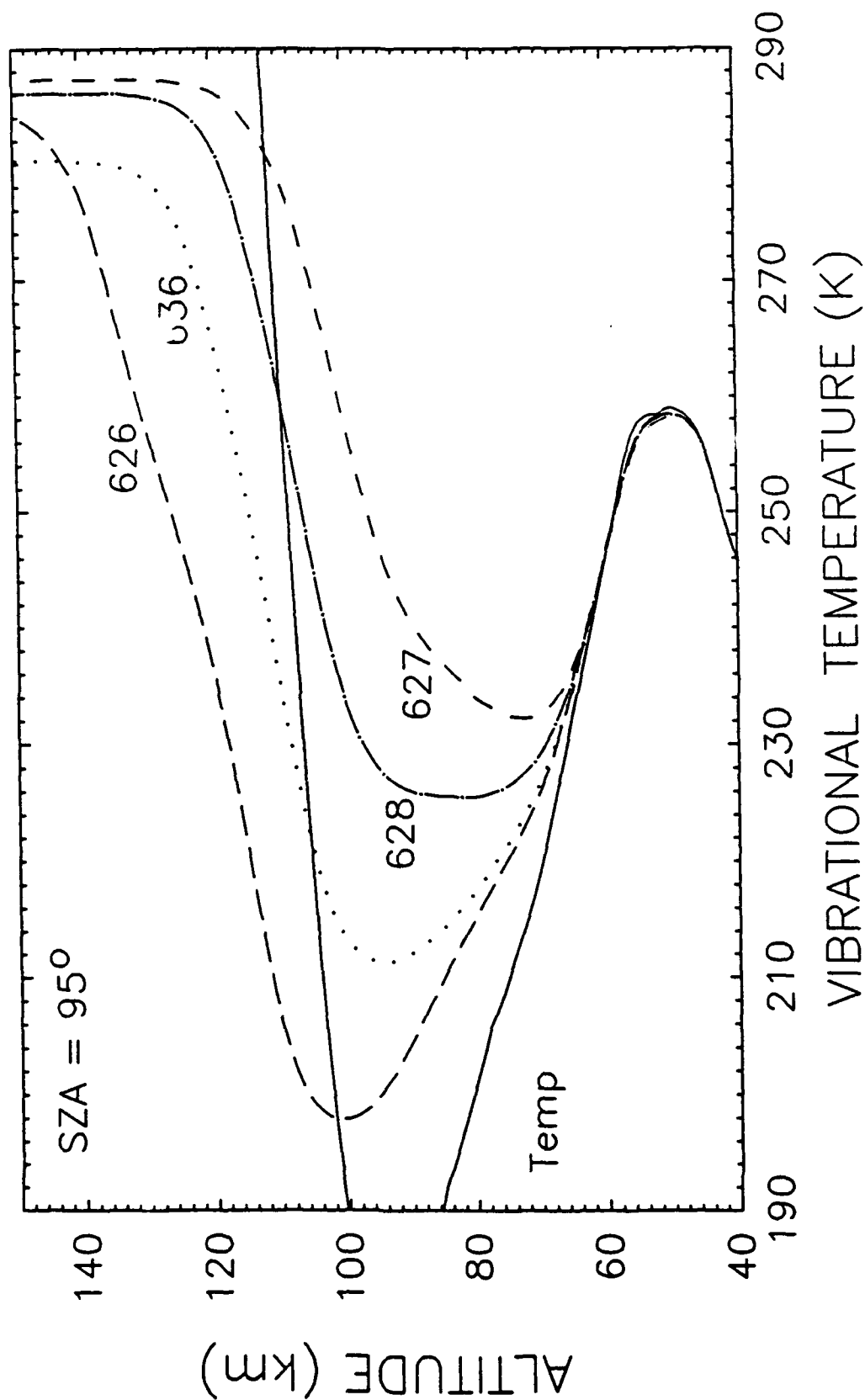
Fig. 9. Vibrational Temperatures for 95° solar zenith angle.
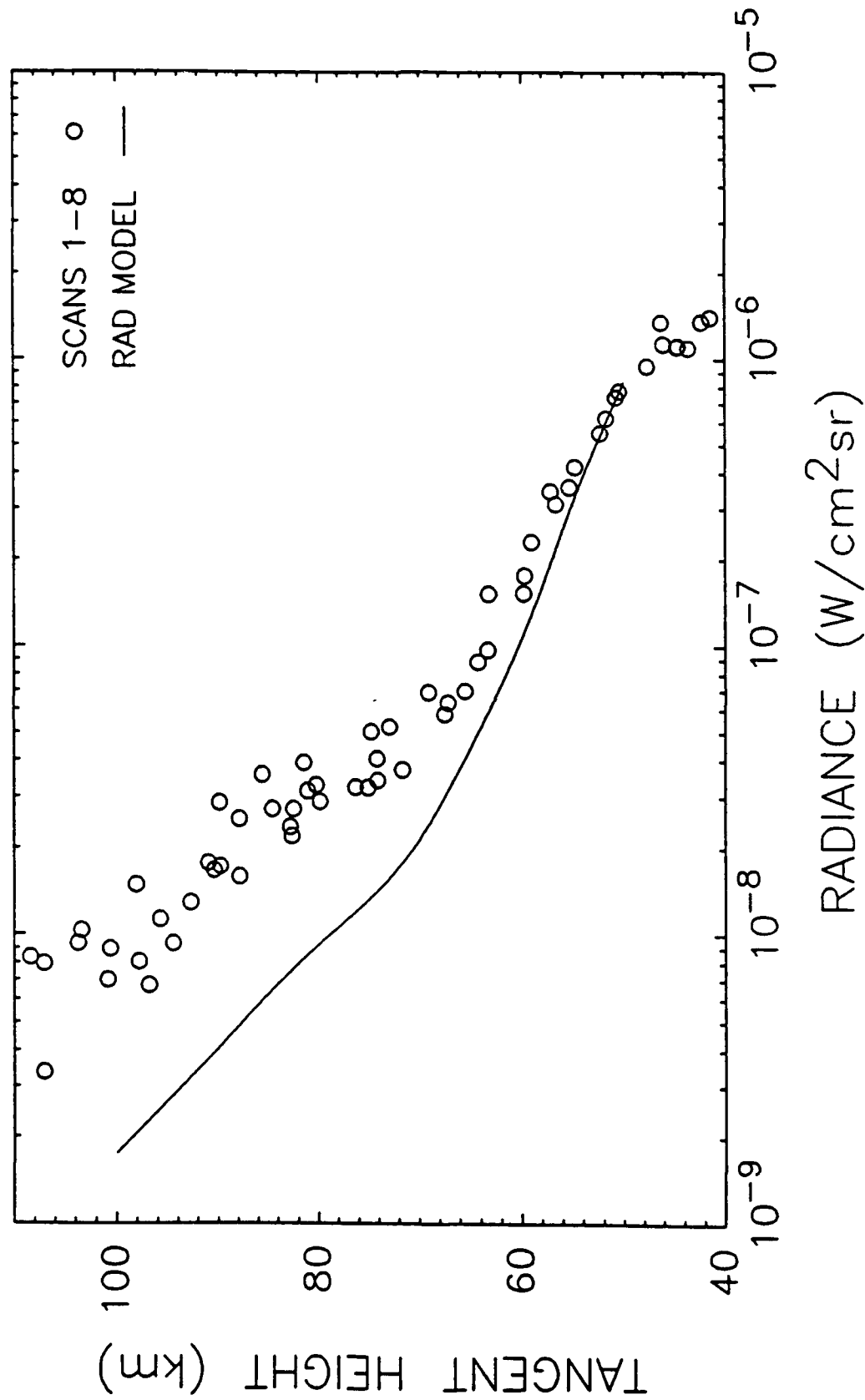
Fig. 10. Total 4.3 micron radiance in a limb view (terminator).

1989-1990 RESEARCH INITIATION PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.

FINAL REPORT

Development and Application of a Dynamo Model
of Electric Fields in the Middle- and Low-
Latitude Ionosphere

Prepared by:	Craig E Rasmussen
Academic Rank:	Assistant Research Scientist
Department and	Atmospheric, Oceanic and Space Sciences
University:	University of Michigan
USAF Researcher:	Nelson C. Maynard Geophysics Laboratory/PHG Hanscom AFB, MA 01731
Date:	2 January, 1991

Development and Application of a Dynamo Model of Electric Fields in the Middle- and Low-Latitude Ionosphere

1. INTRODUCTION AND SUMMARY

This report describes research accomplished under the AFOSR Research Initiation Program for 1989–90. The major thrust of this work was to develop a dynamo model of electric fields in the middle- and low-latitude ionosphere. As part of the testing stage, this model has been used to compliment measurements of a deep ionospheric trough, observed during the great storm of March, 1989. The ionospheric depletions were believed to have been caused by extraordinarily large electric fields which raised the entire F-region ionosphere above the DMSP satellite height. The model has been used to try and understand the origin of these electric fields.

Electric fields in the middle- and low-latitude ionosphere (as well as in the plasmasphere) can be due to either the penetration of magnetospheric electric fields or due to the dynamo action of thermospheric winds. During storms and substorms, it is often difficult to tell which of these two mechanisms is the source for any given set of measurements. However, in many instances, timing, as well as longitudinal variations in these fields, should give an indication as to which of these sources is predominate. It is the ultimate goal of this research to find clear signatures (using the electric field model) of electric fields from each of the two source mechanisms.

The work described in this report is the first step in this research. The penetration of magnetospheric electric fields has been studied during the great storm of March, 1989. In particular, penetrating electric fields were examined to see if they could have been the so ce of the vertical drifts measured by both DMSP F8 and F9 satellites during the storm. This research indicates that the penetrating fields could indeed have been the cause of up d vertical drifts measured ʳear dusk by the F8 satellite, but it is less likely that they are the source of upward drifts measured by the F9 satellite, closer to midnight. Strong coupling between the ionosphere and electric fields of magnetospheric origin is likely. As the height of the F-region ionosphere increases, conductivity gradients should increase near dusk and

this has been found to further enhance vertical drifts and zonal electric fields in this time sector.

From the time that ionospheric depletions were first noticed in the F9 data, the role of the South Atlantic Anamoly (SAA) has been questioned. Strong upward drifts were not seen by F9 until the SAA had rotated eastward of the satellite's orbit. An intriguing question has been whether there is a connection between the timing of the vertical drifts and the location of the SAA? Modeling indicates that E-region conductivity (due to precipitating particles within the SAA) tends to decrease fields of magnetospheric origin within the region of the anomaly. Westward of the SAA, penetrating fields were seen to cause downward drifts as compared to the upward drifts measured by F9. Thus, it is believed that the upward drifts measured by DMSP F9 near midnight are more likely to be of dynamo origin. (Although not discussed in this report, the timing of the F9 observations relative to the storm also supports this conclusion.) It is also noted that the presence of E-region conductivity (inferred from measurements in the region of the SAA [*Batista et al.*, 1991]), would tend to enhance the generation of dynamo electric fields. Further research to understand the effects of the SAA on the production of dynamo electric fields is planned.

Research related to work under this contract has lead to three conference talks at meetings of the American Geophysical Union [*Greenspan et al.*, 1989; *Rasmussen et al.*, 1990; *Greenspan et al.*, 1990] and one publication [*Greenspan et al.*, 1991] is to be submitted to the Journal of Geophysical Research in January, 1991. A draft copy of this paper is attached to the end of this report.

The electric field model is described in the next section. This is followed by an application of the model to the study of the penetration of magnetospheric electric fields during the great storm of March, 1989.

2. MODEL

One of the principal equations for the generation of dynamo electric fields is

$$\mathbf{J} = \sigma \cdot (\mathbf{E} + \mathbf{u} \times \mathbf{B}) \qquad (1)$$

where **J** is current density, **E** is the electric field in Earth's rest frame and σ is the tensor conductivity. The latter term in (1) is the induced emf due to the motion of the neutral gas (**u**) with respect to the geomagnetic field (**B**). This equation explicitly shows that currents flowing in the ionosphere are due to the combined effects of an electric field **E** and the dynamo action of the neutral gas **u** × **B**. In fact, the right-hand-side of (1) is just the electric field in the reference frame of the neutral gas. Should the neutral gas velocity be the same as the plasma convection velocity, $\mathbf{u}_p = \mathbf{E} \times \mathbf{B}/B^2$, no currents are necessary for an electric field to exist in Earth's rest frame. However, in general $\mathbf{u} \neq \mathbf{u}_p$ and currents flow which couple differing regions of Earth's ionosphere. This makes a global solution to (1) necessary, i.e., currents and dynamo winds throughout the global ionosphere must be considered to obtain the electric field at a particular location.

An important assumption is that charge density in the ionosphere is independent of time. Of course whenever temporal changes are important, for instance during substorms when the Birkeland current systems coupling the magnetosphere and the ionosphere are modified, this assumption is invalid. However, temporal changes in charges, currents and electric fields occur extremely rapidly [*Kelley, 1989*] and for time scales of interest, these parameters can be considered to be in pseudo-equilibrium. Thus, the current density is assumed to be divergenceless,

$$\nabla \cdot \mathbf{J} = 0 \qquad (2)$$

and the electric field is assumed to be electrostatic,

$$\mathbf{E} = -\nabla \Phi \qquad (3)$$

The tensor component of conductivity along magnetic field lines is much greater than components perpendicular to the magnetic field. This implies that a natural coordinate system is one that is aligned along the magnetic field direction. A dipole coordinate system is chosen as Earth's magnetic field is closely approximated by a dipole within a few Earth radii. In this system the coordinates are (s, ϕ, q) where s is aligned along the magnetic field (positive northward), ϕ is the azimuthal angle (positive eastward) and q is normal to the other two coordinates. Useful information about the dipole coordinate system includes, the

coordinate transformation:

$$s = \frac{r_0^2}{r^2} \cos \theta \tag{4}$$

$$q = \frac{r_0}{r} \sin^2 \theta \tag{5}$$

the unit vectors:

$$\hat{\mathbf{e}}_s = -\sin(\chi)\hat{\mathbf{e}}_r - \cos(\chi)\hat{\mathbf{e}}_\theta \tag{6}$$

$$\hat{\mathbf{e}}_q = -\cos(\chi)\hat{\mathbf{e}}_r + \sin(\chi)\hat{\mathbf{e}}_\theta \tag{7}$$

and the magnitudes of the coordinate base vectors:

$$h_s = \frac{r^3}{r_0^2\sqrt{1 + 3\cos^2 \theta}} \tag{8}$$

$$h_\phi = r \sin \theta \tag{9}$$

$$h_q = \frac{r^2}{r_0 \sin \theta \sqrt{1 + 3\cos^2 \theta}} \tag{10}$$

where r is the radius from the dipole center, r_0 is a reference altitude (90 km), θ is colatitude, $\hat{\mathbf{e}}_r$ is a unit vector along r, $\hat{\mathbf{e}}_\theta$ is unit vector along θ (positive southward) and χ is the dip angle defined by

$$\cos(\chi) = \frac{\sin \theta}{\sqrt{1 + 3\cos^2 \theta}} \tag{11}$$

$$\sin(\chi) = \frac{2\cos \theta}{\sqrt{1 + 3\cos^2 \theta}} \tag{12}$$

$$\tan(\chi) = 2\cot(\theta) \tag{13}$$

Another parameter is frequently used:

$$L = \frac{1}{q} = \frac{r}{r_0 \sin^2 \theta} \tag{14}$$

Given the above information, (1) can be substituted into (2), to obtain

$$a\frac{\partial^2 \Phi}{\partial q^2} + b\frac{\partial^2 \Phi}{\partial \phi^2} + d\frac{\partial \Phi}{\partial q} + e\frac{\partial \Phi}{\partial \phi} + f = \frac{\partial}{\partial s}(h_\phi h_q j_s) \tag{15}$$

where j_s is the component of current density along s and where

$$a = \frac{h_s h_\phi}{h_q}\sigma_1 = h_a \sigma_1 \tag{16}$$

$$b = \frac{h_s h_q}{h_\phi}\sigma_1 = h_b \sigma_1 \tag{17}$$

$$d = \frac{\partial h_a \sigma_1}{\partial q} - \frac{\partial h_s \sigma_2}{\partial \phi} \tag{18}$$

$$e = \frac{\partial h_s \sigma_2}{\partial q} + \frac{\partial h_b \sigma_1}{\partial \phi} \tag{19}$$

$$f = \frac{\partial}{\partial q}\left[h_s h_\phi B(\sigma_1 u_\phi - \sigma_2 u_\theta \sin \chi)\right] - \frac{\partial}{\partial \phi}\left[h_s h_q B(\sigma_2 u_\phi + \sigma_1 u_\theta \sin \chi)\right] \tag{20}$$

with σ_1 representing Pedersen conductivity, σ_2 Hall conductivity and u_θ and u_ϕ are neutral winds in the θ and ϕ directions, respectively.

Equation (15) is second-order partial differential equation in three dimensions. Considerable simplification is possible if one takes into account the very large conductivity in the magnetic field direction. This implies that parallel electric fields are very small and that almost all potential drops occur in a direction perpendicular to field lines. Thus, it is further assumed that the potential Φ is constant along a field line allowing (15) to be integrated from the southern hemisphere to the northern hemisphere. This reduces the dimensions of the system by 1. Integrating (15) gives

$$A\frac{\partial^2 \Phi}{\partial q^2} + B\frac{\partial^2 \Phi}{\partial \phi^2} + D\frac{\partial \Phi}{\partial q} + E\frac{\partial \Phi}{\partial \phi} + F = 0 \tag{21}$$

where

$$A = \int_{s_S}^{s_N} a\, ds \tag{22}$$

$$B = \int_{s_S}^{s_N} b\, ds \tag{23}$$

$$D = \int_{s_S}^{s_N} d\, ds + \left[\frac{\sigma_1 h_\phi \tan \theta}{2}\right]_{s_N} + \left[\frac{\sigma_1 h_\phi \tan \theta}{2}\right]_{s_S} \tag{24}$$

$$E = \int_{s_S}^{s_N} e\, ds + \left[\frac{\sigma_2 h_q \tan\theta}{2}\right]_{s_N} + \left[\frac{\sigma_2 h_q \tan\theta}{2}\right]_{s_S} \qquad (25)$$

$$F = \int_{s_S}^{s_N} f\, ds + \left[\frac{h_\phi h_q B \tan\theta}{2}(\sigma_1 u_\phi - \sigma_2 u_\theta \sin\chi)\right]_{s_N}$$
$$+ \left[\frac{h_\phi h_q B \tan\theta}{2}(\sigma_1 u_\phi - \sigma_2 u_\theta \sin\chi)\right]_{s_S} - h_\phi h_q j_{ext} \qquad (26)$$

Note the additional terms in the equations for the coefficients D, E, and F. These extra terms are necessary to ensure that no current leaks out of the lower ionospheric boundaries of the system (s_N and s_S). They arise from setting $j_r = 0$ at the two boundaries [*Takeda and Maeda*, 1980]. There is also an additional term in (26), j_{ext}, which allows for the addition of external currents, for instance, the region I and II Birkeland currents.

Equation (21) is the two-dimensional dynamo equation which must be solved numerically. This equation was solved on the two-dimensional grid, $\phi = 0 : 2\pi$ and $q = 0 : 1$. Boundary conditions are necessary at $q = 0$ (the pole) and at $q = 1$ (the equator). At the pole the potential was set to a constant value while a derivative boundary condition

$$\frac{\partial\Phi}{\partial q} = -\frac{\sigma_2}{\sigma_1}\frac{\partial\Phi}{\partial\phi} - r_0 B_0 u_\phi \qquad (27)$$

was assumed at the equator to ensure that $j_r(q = 1) = 0$, where B_0 is the value of the magnetic field at the equator. Equation (21) was solved using line SOR on a 48×90 grid (48 grid points in the east-west direction).

The dipole coordinate system employed allows many factors which are important in the equatorial ionosphere to be considered. These include: (1) the coupling of the E and F region ionospheres; (2) horizontal magnetic field lines in the equatorial region and (3) conductivity and wind asymmetries between the northern and southern ionospheres. These north-south asymmetries imply a parallel current coupling the two hemispheres and this coupling is self-consistently included in the model.

2.1. *Results of the Model*

Given a model of conductivities and the neutral wind (along with appropriate boundary conditions), these equations can be solved for the electrostatic potential Φ. An example of a

solution to (21) is shown in Figure 1, where contours of constant electrostatic potential are plotted at a reference altitude of 300 km. The thermospheric wind pattern for this solution was obtained from NCAR TGCM simulations of the September 18 Equinox Transition Study (ETS) period [*Crowley et al.*, 1989]. The universal time is 0000 UT and corresponds to a magnetically quiet period at the beginning of the study. The conductivity model used is idealized and does not take into account variations in conductivity due to variations in the solar zenith angle. Therefore, it is stressed that the results shown in Figure 1 are only preliminary. However, a comparison with the results of a similar model [*Richmond and Roble*, 1987] and with an empirical model [*Richmond et al.*, 1980], indicates that the numerical model described above is apparently functioning correctly.

However, a careful examination of the flow of current indicates that current is not fully conserved in this numerical model. This can be seen in Figure 2 where horizontal current vectors are plotted as a function of local time and latitude. If the meridional component of the top row of vectors (about 40° latitude) is considered, it is apparent that more current leaves the poleward boundary in the dusk sector than enters in the dawn sector. As there were no dynamo winds or external sources of current applied in this test case, this implies that there is a net loss of current due to numerical errors. After careful consideration of the numerical method and in consultation with B. van Leer (a numerical analyst at the University of Michigan), it was found that the numerical solution to (21) is not written in conservative form and so current is not conserved (to second order), even though the numerical solution to (21) is accurate to second order. A new numerical algorithm has been designed and is currently being coded.

Because of the lack of current conservation in the present numerical method, the solutions described below are expected to give qualitatively the correct results, but are not precisely quantitatively correct.

3. APPLICATION

The electric field model described above was used to perform several numerical experiments to try to pinpoint possible mechanisms for abnormally large upward drifts (eastward electric fields) measured by DMSP F8 and F9 during the great storm of March. 1989. Part

of my research activity at the Geophysics Laboratory was spend studying electric fields and their effects on the equatorial ionosphere during the storm. A paper, to be submitted to the Journal of Geophysical Research describing part of this research, is attached to the end of this report. Other results were presented at the Fall meeting of the American Geophysical Society. These results are described below.

Two possible mechanisms for the abnormally large electric fields measured at dusk and at 2200 LT are likely: (1) the direct penetration of magnetospheric fields and (2) dynamo fields created by disturbed thermospheric winds. The electric field model was used to examine the effects of the direct penetration of magnetospheric fields to see if these fields could have been responsible for the measured upward drifts. Electric fields of magnetospheric origin were introduced as a boundary condition at 45° latitude and the model was run until convergence. This allowed a self-consistent solution to the penetration of magnetospheric fields to be obtained. No fields of dynamo origin were considered as the goal was to examine the effects of the penetrating fields only. Four questions were specifically addressed as detailed below.

3.1. *Expansion of the Auroral Zone*

During the storm, auroral precipitation was seen at abnormally low latitudes ($\sim 45°$). Associated with the precipitation were large westward drifts in the evening sector at similar latitudes. These drifts were presumably part of the evening convection cell that had expanded to low latitudes. The question addressed here is whether an appreciable portion of the Region I current system (associated with the normal dusk and dawn convection cells) closes via the equatorial ionosphere. If so, this current could produce low-latitude electric fields that are consistent with the upward drifts measured by the DMSP satellites.

As can be seen in Figure 2, the answer is yes. Equatorward current is seen to flow out of the morning midlatitude region until it nearly reaches the equator. Here the current is diverted eastward as it forms part of the equatorial electrojet until it eventually returns poleward in the dusk sector. This effect is not surprising and has been modeled before [e.g., *Nopper and Carovillano,* 1978], however, it was found that the percentage of Region I current diverted to the equatorial region is enhanced, simply by the proximity of the auroral zone to the highly conducting equatorial region. No increase in Region I current strength is needed to get increased fields at low latitudes, just an expansion of the Region I current system to

low latitudes. Any increase in Region I current strength would further enhance the electric fields at low latitudes. Thus, enhanced low-latitude fields are expected during the storm (at least until shielding has a chance to occur). Whether these fields can be expected to have the correct direction in the evening sector is now examined.

3.2. *Rotation of Northward Electric Fields near the Equator*

Poleward (and upward) electric fields near the equator (responsible for the observed westward drifts believed to be associated with the evening convection cell), require that there be a Pedersen current flowing toward the poles on either side of the equator. This implies a net divergence of current from the equatorial region unless there is a corresponding convergence of current from another source. If the balancing current is a Hall current flowing equatorward, then an eastward electric field must be present to drive the Hall current. This eastward electric field component could then be the source of the measured vertical drifts and would be responsible for an eastward rotation of the northward field near the equator.

This rotation can be seen in Figure 3 where the Pedersen current density is represented by vectors. Pedersen current is aligned with the electric field so the direction of these vectors represents the electric field direction as well. Note the significant eastward component of Pedersen current in the afternoon and early evening sectors below 5° latitude. By comparing current vectors in this region with those at slightly higher latitudes, it is clear that there is an eastward rotation of the current (and electric field) near the equator.

As there is a net meridional current (Pedersen) flowing poleward from the equator, clearly there is a divergence of Pedersen current as discussed above. Thus, a balancing convergence of current due to a meridional Hall current must be present. This equatorward Hall current can be seen in Figure 4, where Hall current vectors are plotted. The Hall current contains a significant equatorward component in the afternoon and early evening sectors below 5° latitude. This equatorward Hall current balances the poleward Pedersen current (see the net current shown in Figure 2) and is driven by an eastward electric field. This demonstrates that there is a linkage between westward drifts at middle to high latitudes and an upward drift at low latitudes.

Could this mechanism be the source for the upward drifts measured by the DMSP satel-

lites during the great storm of March, 1989? Apparently not, because a significant eastward rotation of Pederson current is only seen at latitudes below about 5°. This latitude range corresponds to altitudes below 150 km. As the altitude of the DMSP satellites is approximately 840 km, the satellites only make measurements at latitudes (corresponding to the mappings in Figures 3 and 4) above roughly 18°. Thus, some mechanism, other than an eastward rotation of a poleward electric field (to preserve current continuity) must be responsible for the observations of abnormally large upward drifts in the evening sector.

3.3. *Effects of Reduced F-Region Conductivity*

An interesting question arises when one considers feedback. Upward drifts lead to density depletions below the F-region peak and thus to decreases in F-region conductivity. Do decreases in conductivity affect electric fields in a way which could create a positive feedback mechanism to further enhance the upward drifts?

To ascertain this possiblity, conductivity was reduced in the nighttime equatorial ionosphere within 20° of the equator. The results of this numerical experiment are shown in Figure 5, where surfaces of constant potential are plotted in 5a for the normal case (some nighttime F-region conductivity) and in 5b for the depleted case. These potential contours are also flow lines as denoted by the arrows which show the direction of the flow. A comparison of these two figures, shows that enhanced upward (and poleward) drifts occur as a result of the modeled conductivity decrease near 2200 LT (Figure 5b). Conductivity gr dients in this region of the dusk ionosphere (0–20° latitude), enhance electric fields near the gradients and produce a reversal in the ion flows with upward (poleward) flows to the dayside of the reversal and downward (equatorward) flows to the nightside. Thus, this mechanism, while increasing the likelihood of upward drifts in the early evening sector, would tend to decrease the length in local time that upward drifts take place.

3.4. *Effects of E-Region Conductivity Enhancements*

The strong upward drifts seen by DMSP F9 were measured on the first two orbits after it left the region of the South Atlantic Anomaly (SAA). Is this timing coincidental or does the SAA play a role in the F9 measurements? Increased E-region conductivity is inferred from ion density measurements in the SAA [*Batista et al.*, 1991] and it would be expected

that this conductivity would have an effect on the electric fields and drifts in this region of the ionosphere.

The effect of increased conductivity can be seen in Figure 6 where lines of constant electric potential are plotted. The most obvious effect of the increase in conductivity is to reduce electric fields within the region of enhanced conductivity (shown by cross-hatched lines), therebye diverting flow around the region. While a conductivity enhancement may explain the lack of upward drifts measured by DMSP F9 while it was within the SAA, the modeled direction of the flow is in the wrong sense, at longitudes westward of the SAA. Therefore, it is unlikely that the SAA, in conjunction with penetrating magnetospheric fields, played an important role in the measured upward drifts by DMSP F9.

However, increased conductivity in the SAA would act to enhance the generation of dynamo electric fields. Perturbed thermospheric winds (a source of dynamo electric fields in the equatorial ionosphere [*Blanc and Richmond*, 1980]) are a likely consequence of increased heating in the thermosphere associated with the ongoing magnetic storm. Further numerical experiments are planned to study the potential importance of thermospheric winds in the region of the SAA during the storm.

4. CONCLUSIONS

The electric field model has been applied to a study of the penetration of magnetospheric electric fields during the great storm of March, 1989. Preliminary results suggest that: (1) The orientation of magnetic fields near the equator does indeed cause an eastward rotation of poleward electric fields (as originally thought). However, this rotation is confined to very low latitudes and cannot be responsible for the verticle drifts measured by the DMSP satellites. (2) Conductivity gradients near the dusk terminator cause a velocity shear and act to enhance verticle drifts. This may have been responsible for the verticle drifts measured near dusk by DMSP F8 but not for the verticle drifts near midnight measured by F9. (3) E-region conductivity enhancements in the region of the South Atlantic Anomaly make it unlikely that F9 drifts measurements were caused by the penetration of magnetospheric fields. Thus, dynamo electric fields may have been responsible for the upward drifts seen by F9. The generation of dynamo electric fields will be increased by conductivity enhancements in the SAA.

REFERENCES

Batista, I. S., E. R. de Paula, M. A. Abdu, and N. B. Trivedi, Ionospheric effects of the 13 March 1989 magnetic storm at low and equatorial latitudes, submitted to *J. Geophys. Res.*, January, 1991.

Blanc, M., and A. D. Richmond, The ionospheric disturbance dynamo, *J. Geophys. Res.*, *85*, 1669, 1980.

Crowley, G., B.A. Emery, R.G. Roble, H.C. Carlson, Jr., and D.J. Knipp, Thermospheric dynamics during September 18–19, 1984: I. Model simulations, *J. Geophys. Res.*, *94*, 16,925–16,944, 1989.

Greenspan, M. E., W. J. Burke, and C. E. Rasmussen, Ion density depletions at 840 km over the South Atlantic anomaly during the magnetic storm of March, 1989, AGU Fall Meeting, San Francisco, California; *EOS Trans., AGU, 70*, 1259, 1989.

Greenspan, M. E., C. E. Rasmussen, W. J. Burke, and M. A. Abdu, Interpretation of large storm-induced holes in the equatorial ionosphere, AGU Fall Meeting, San Francisco, California; *EOS Trans., AGU, 71*, 1489, 1990.

Greenspan, M. E., C. E. Rasmussen, W. J. Burke, and M. A. Abdu, Equatorial density depletions observed at 840 km during the great magnetic storm of March, 1989, submitted to *J. Geophys. Res.*, January, 1991.

Kelley, M. C., *The Earth's Ionosphere: Plasma Physics and Electrodynamics*, Academic, San Diego, Calif., 1989.

Nopper, R. W., and R. L. Carovillano, Polar equatorial coupling during magnetically active periods, *Geophys. Res. Lett.*, *5*, 699, 1978.

Rasmussen, C. E., G. Crowley, and C. Coray, Disturbance-dynamo electric fields during the equinox transition study of September, 17–24, 1984, AGU Fall Meeting, San Francisco, California; *EOS Trans., AGU, 71*, 1490, 1990.

Richmond, A. D., and R. G. Roble, Electrodynamic effects of thermospheric winds from the NCAR thermospheric general circulation model, *J. Geophys. Res.*, *92*, 12,365–12,376, 1987.

Richmond, A. D., M. Blanc, B. A. Emery, R. H. Wand, B. G. Fejer, R. F. Woodman, S. Ganguly, P. Amayenc, R. A. Behnke, C. Calderon, and J. V. Evans, An empirical model

of quiet-day ionospheric electric fields at middle and low latitudes, *J. Geophys. Res.*, *85*, 4658–4664, 1980.

Takeda, M. and H. Maeda, Three-dimensional structure of ionospheric currents 1. Currents caused by diurnal tidal winds, *J. Geophys. Res.*, *85*, 6895, 1980.
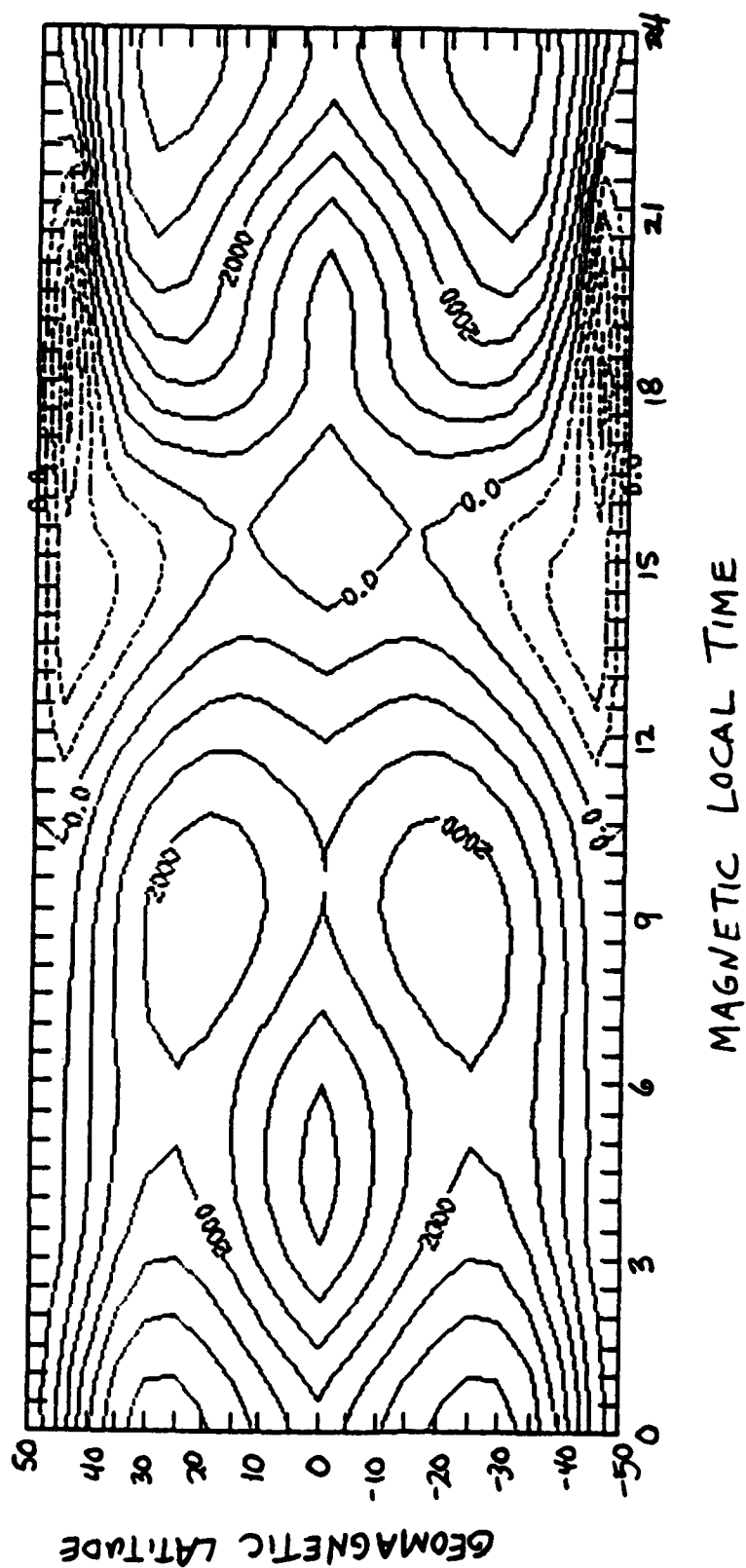
Fig. 1. Contours of constant electrostatic potential at a reference altitude of 300 km. The electric fields are of dynamo origin. The contour interval is 0.5 kV.
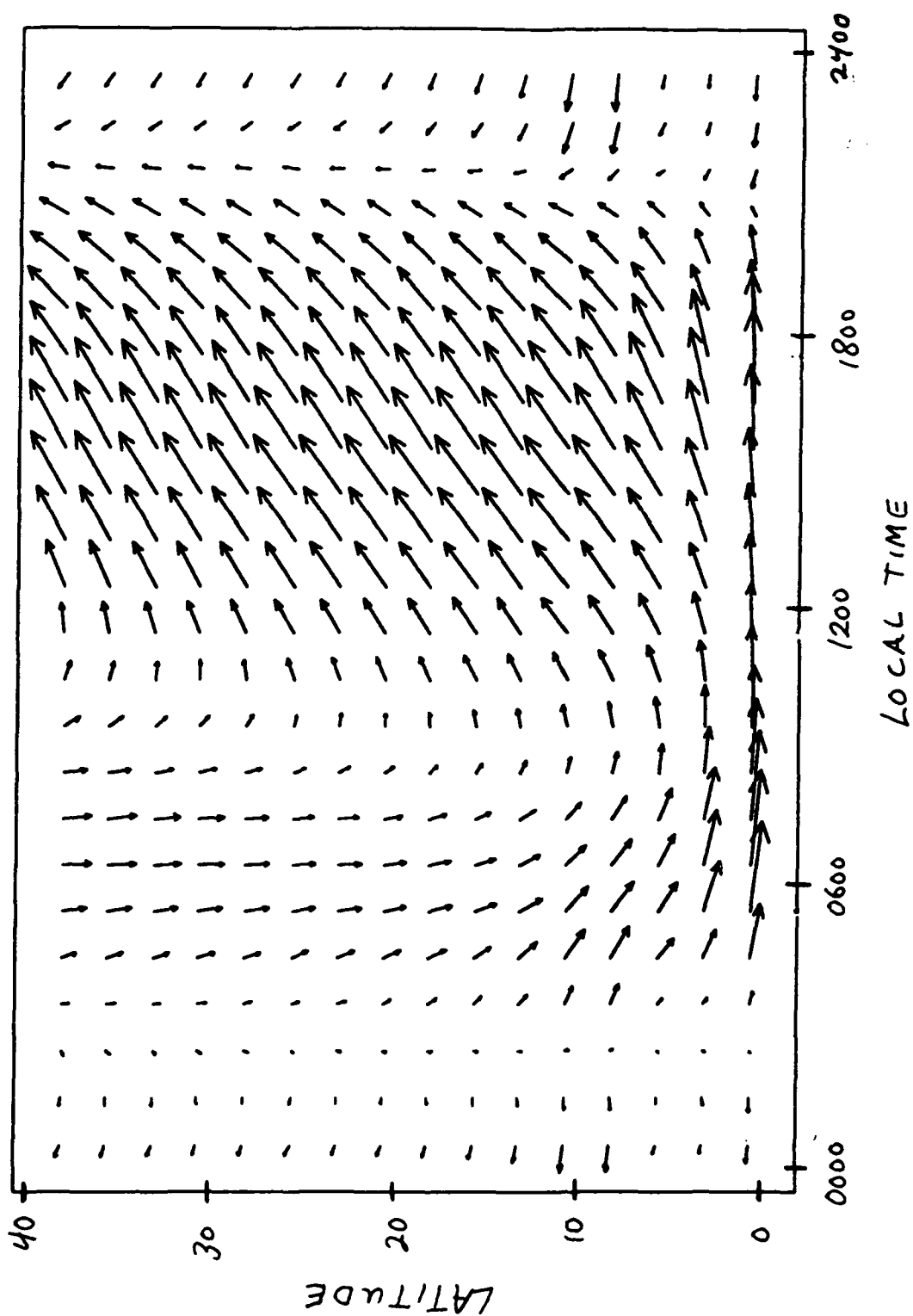
Fig. 2. Height integrated total current density. The magnitude of the largest vectors is approximately 0.1 A/m.
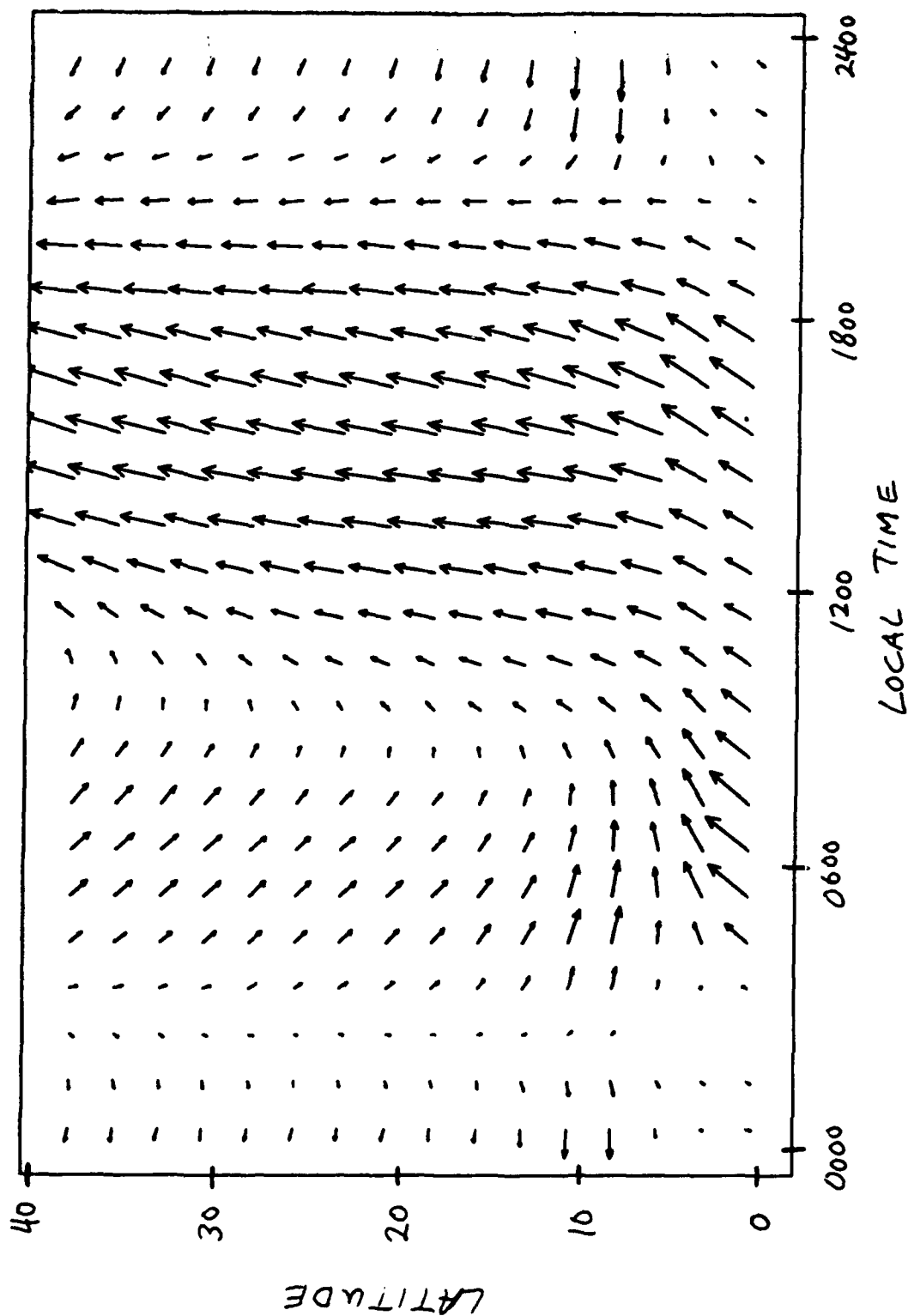
Fig. 3. Height integrated Pedersen current density. The magnitude of the largest vectors is approximately 0.1 A/m.
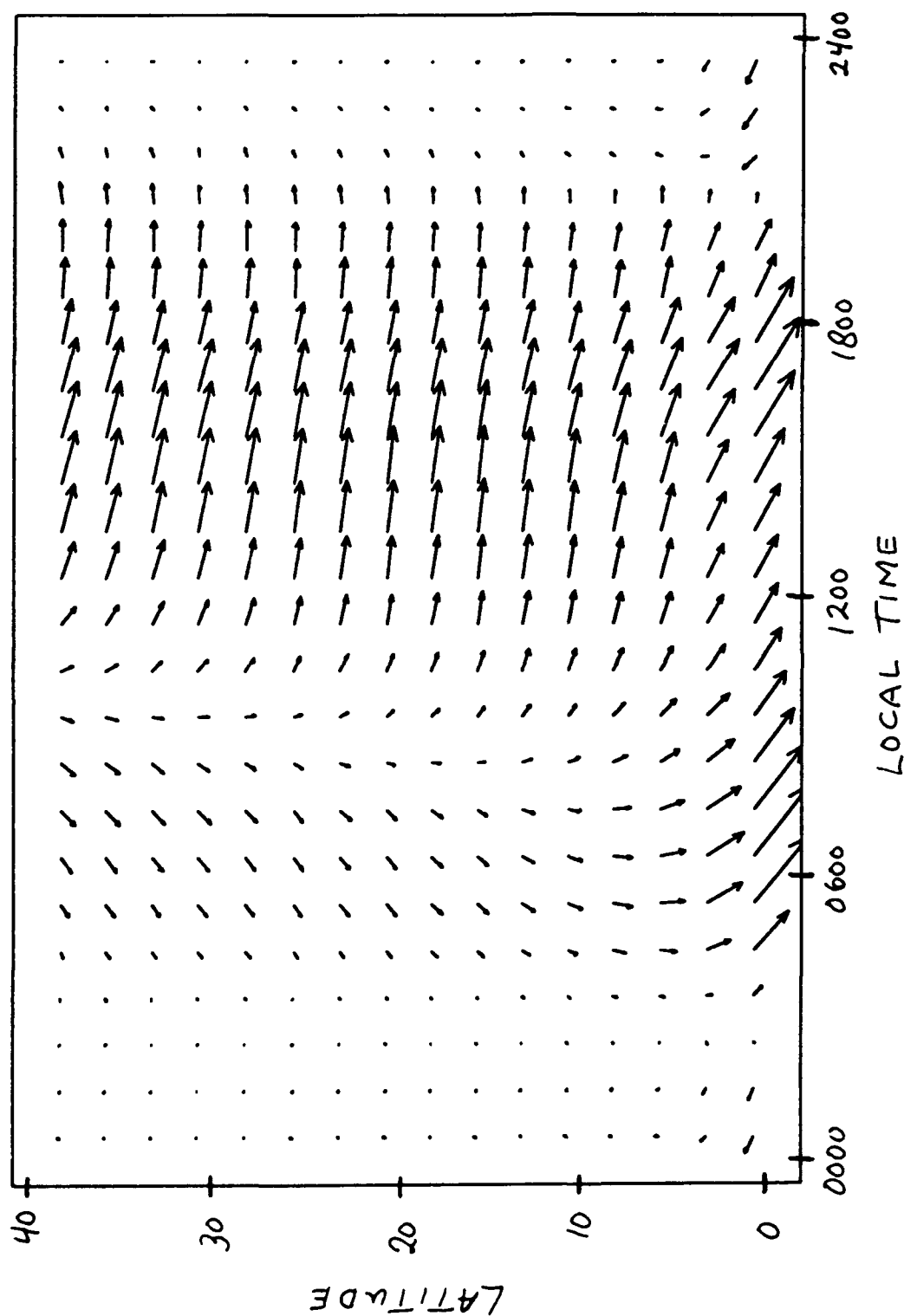
Fig. 4. Height integrated Hall current density. The magnitude of the largest vectors is approximately 0.1 A/m.
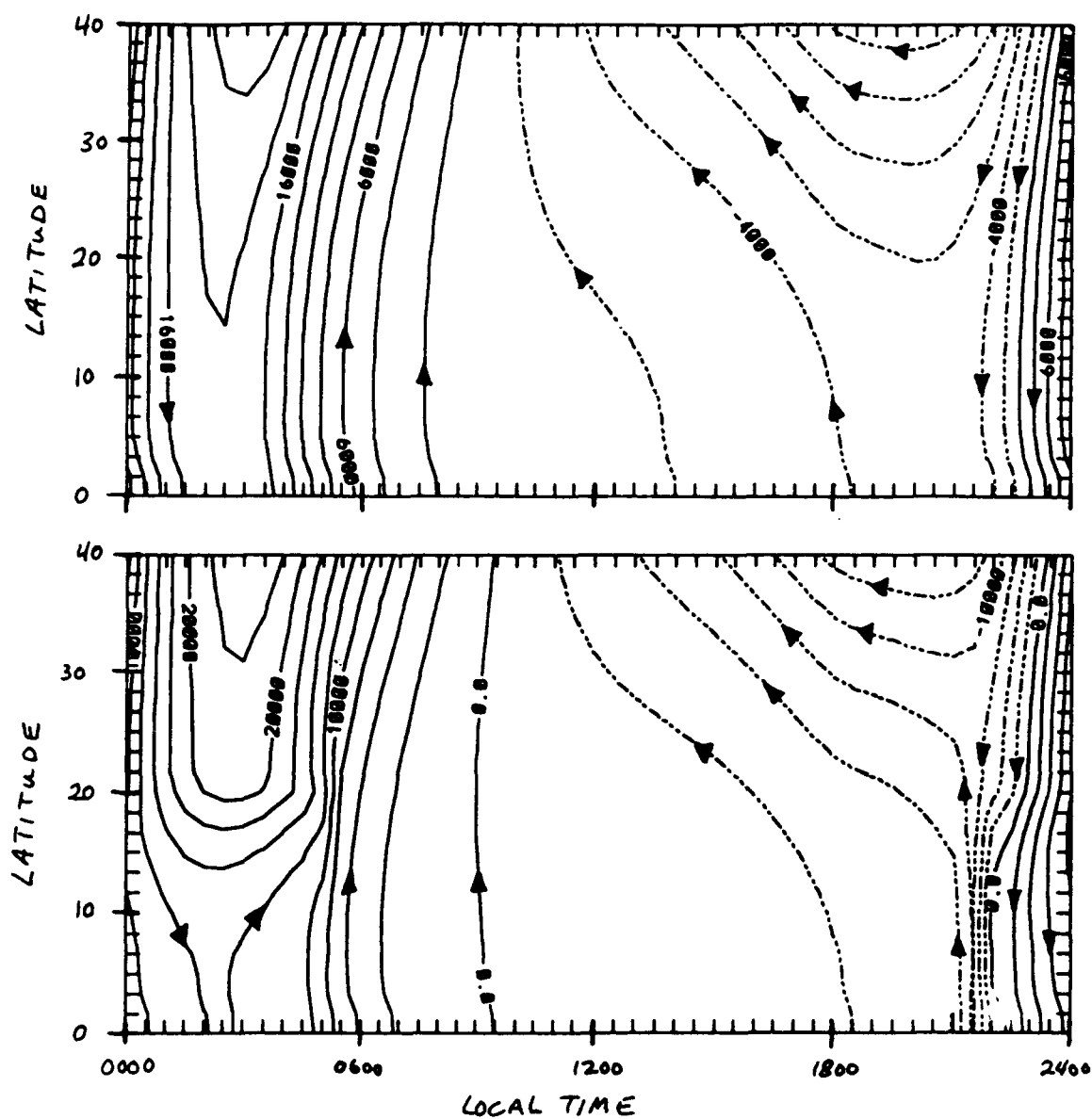
Fig. 5. Contours of constant electrostatic potential at a reference altitude of 90 km for (a) regular conductivity distribution and (b) reduced F-region conductivity in the nighttime equatorial ionosphere. The contour interval is 2.5 kV.
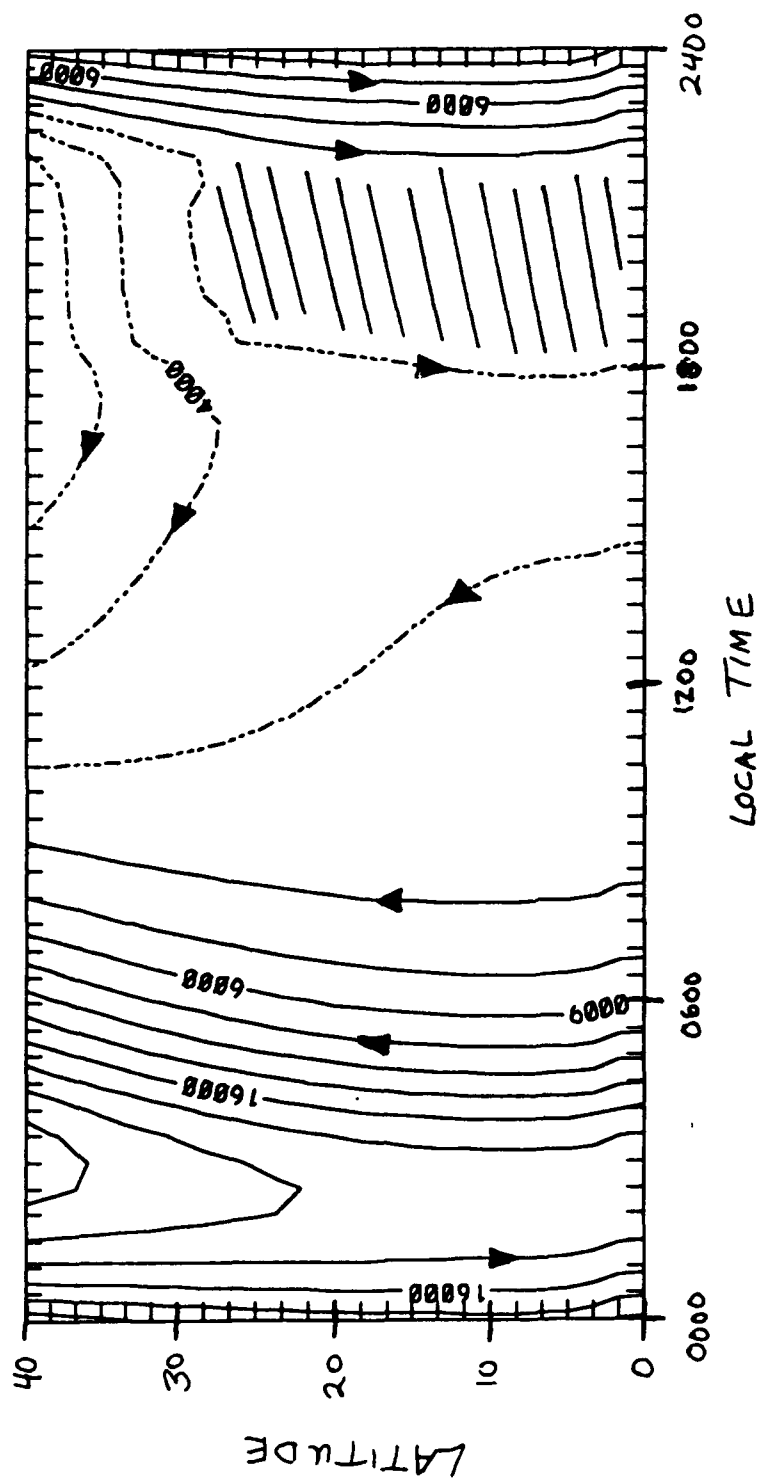
Fig. 6. Contours of constant electrostatic potential at a reference altitude of 90 km for a conductivity enhancement in the SAA (the cross-hatched region). The contour interval is 2.5 kV.

FINAL REPORT - RESEARCH INITIATION PROGRAM (RIP)

Sponsored by the
AIR FORCE OFFICE OF SCIENTIFIC RESEARCH
Conducted by
Universal Energy Systems, Inc.

SYNTHESIS OF TETRANITROHOMOCUBANE

Principal
Investigator: Lynn Maruyama Kirms, Ph.D.

Academic Rank: Assistant Professor

Department and
University: Chemistry Department, Southern Oregon State College
                     Ashland, Oregon 97520

Date: December 31, 1990

Contract No: F49620-88-C-0053/SB5881-0378

I.   BACKGROUND:

The RIP proposal submitted a year ago was a request for funds to work on the synthesis of tetranitrohomocubane. A portion of the proposed synthesis was based on methodology developed for the synthesis of pentanitrobishomocubane, a project the author was involved in during the summer of 1989 as an SFRP fellow at Edwards AFB.

The tetranitrohomocubane system is of interest because its carbon skeleton is more strained than that of the bishomocubane molecule (a comparison is being made to the pentacyclo[$5.3.0.0^{2,5}.0^{3,9}.0^{4,8}$]decane form of bishomocubane) as homocubane lacks one of the methylene groups. Force field calculations indicate that the strain energy of homocubane is on the order of 125 kcal/mol[1], while the strain energy of pentacyclo[$5.3.0.0^{2,5}.0^{3,9}.0^{4,8}$]decane is approximately 83 kcal/mol[2]. The strain in a cage system is a factor is considering its potential for use as a propellant or as an explosive.

Scheme 1 shows the proposed synthesis of the cage compound leading to tetranitrohomocubane as outlined in the author's RIP proposal. The key precursor in this synthesis is the diol 1, which could be prepared in five steps[3,4] starting from p-benzoquinone and cyclopentadiene as outlined in Scheme 2.   Oxidation of diol 1 would give the diketone 2. This diketone, when treated with MCPBA, would yield the epoxide 3 which could undergo a Favorskii type ring contraction to afford 4. The Favorskii step would accomplish two goals: (i) the ring would be contracted and ready to undergo the photochemical [2+2] cyclization to the homocubane skeleton and (ii) a carbomethoxy group would be strategically placed for further elaboration to a nitro group. After the photochemical cyclization, the carbomethoxy moieties and the ketone of 5 would be transformed into nitro and geminal di-nitro species respectively by use of established methodologies[3] to afford the final product, 6.

Scheme 1

MeO₂C — oxidation → 2 — MCPBA → 3

3 — NaOH, MeOH / Favorskii Ring Contraction → 4

4 — [2+2] → 5

5 → Tetranitrohomocubane 6

Scheme 2

7 + (quinone) — Diels Alder → 9 — H₂O₂ / Na₂CO₃ → 10

NaOH, MeOH / Favorskii ring contraction → 11 — NaBH₄ / Ce(III) / MeOH → [intermediate] — [3,3] → 12

12 — SeO₂ / allylic oxidation → 1

44-3
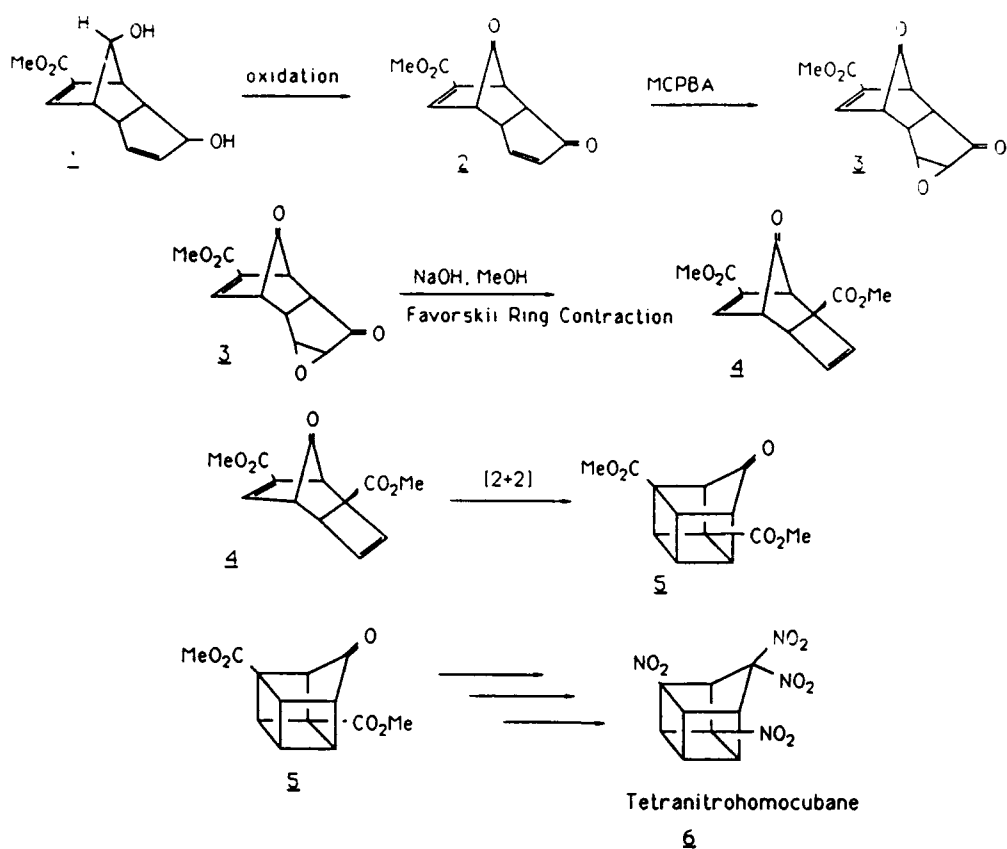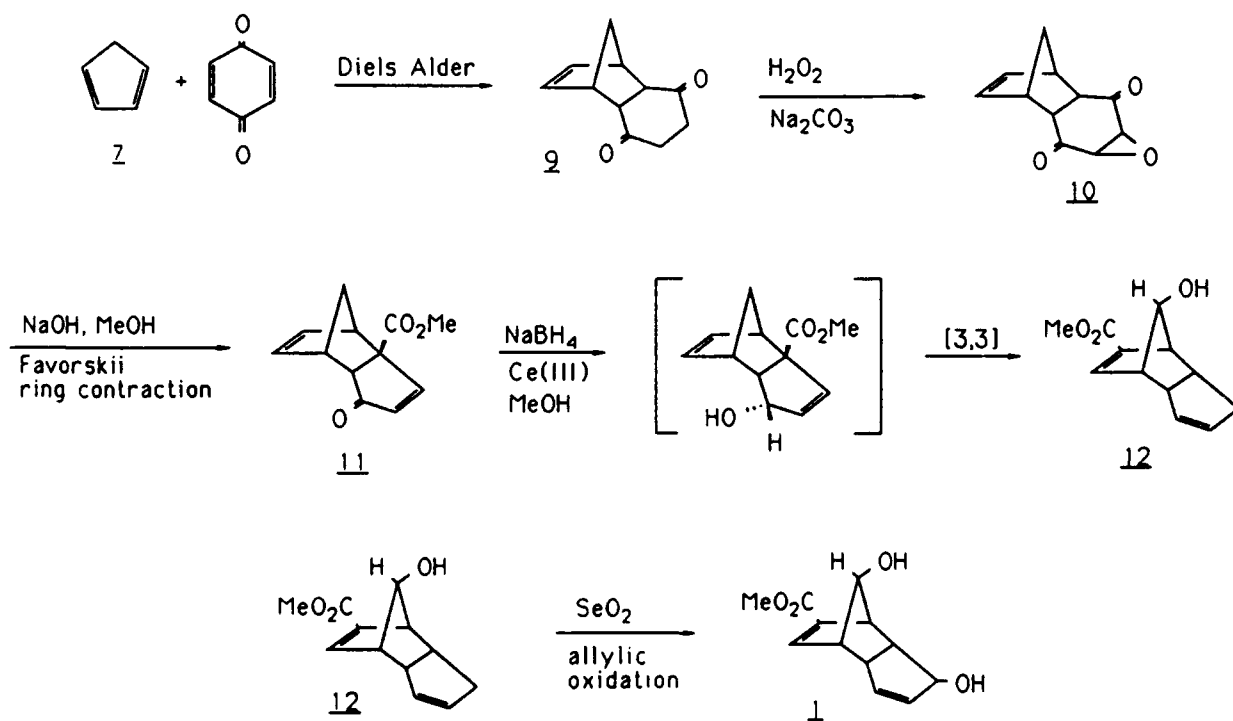
II. SYNTHETIC INVESTIGATIONS:

In synthesizing diol 1 to have enough material to work with, we encountered much difficulty with the Favorskii ring contraction (the conversion of 10 to 11, see Scheme 2). The yield is reported to range from 55-70%,[5,6] but despite our best efforts and many attempts to improve the reaction, we were only able to realize a 25% yield at best.

The oxidation of diol 1 to the diketone 2 was the focus of other efforts. Treatment of 1 with PCC had resulted in oxidation of the alcohol on the 5-membered ring but the bridge alcohol remained untouched. This could be rationalized by considering the strain involved in placing an sp^2 center on the bridge as opposed to making an enone in the 5-membered ring. A Jones oxidation of 1 yielded the same results as the PCC oxidation in very low yield (6%). None of the desired product was isolated, and the only other product from the reaction appeared to have lost the methyl group of the ester due to the strongly acidic conditions of the Jones reagent. A milder Swern oxidation was attempted on 1; this gave an intractable mixture of products.

At this juncture it was decided to try to oxidize the bridge alcohol before the SeO_2 mediated allylic oxidation that yielded the diol (Scheme 3). In this way, we could eliminate any interference that the alcohol on the 5-membered ring might contribute to the overall oxidation. The allylic oxidation of the 5-membered ring and subsequent transformation to a carbonyl could take place at a later time. A Jones and a Collins oxidation were carried out on compound 12 and yielded some interesting results. We found that the desired bridge ketone 14 was formed in both cases, however, this ketone rearranged in a [3,3] sigmatropic fashion back to compound 12. For the Jones oxidation, rearrangement occurred during the course of the reaction, but in the Collins oxidation, the rearrangement occurred in the workup. Upon searching the literature, this occurrence was found to be well documented.[7,8]

Scheme 3

12   14   11
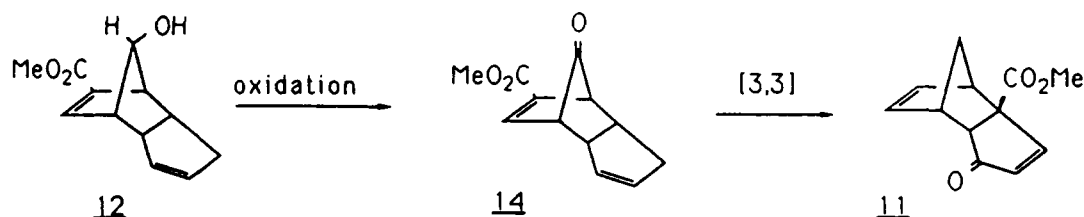
It was thought that suitable reaction conditions might be found wherein compound 14 would be stable. Indeed, Klunder et.al.[7] claimed that subjecting 12 to flash vacuum thermolysis would give the desired 14 but details were not given. We tried a sealed tube reaction on 12 (175 C, 2 days) which yielded five products without consuming all of the starting material. The forcing conditions necessary to go against the thermodynamics of the reaction (from an enone system in 12 to a bridged ketone in 14) unfortunately result in decomposition of the reaction mixture. This approach was abandoned.

A strategy currently being pursued is the synthesis of compound 16 as shown in Scheme 4. The bridge alcohol of 12 is protected as the t-butyldimethylsilylether and carried through in protected form to the cage compound 16. Once 16 is deprotected, it is hoped that the resulting bridge alcohol can then be oxidized to give the desired ketone, compound 5 (Scheme 1). As noted earlier, the carbomethoxy and ketone groups of 5 could then be transformed into nitro groups to give the tetranitrohomocubane. Work on this project will continue along these lines.

Scheme 4

12   15   16
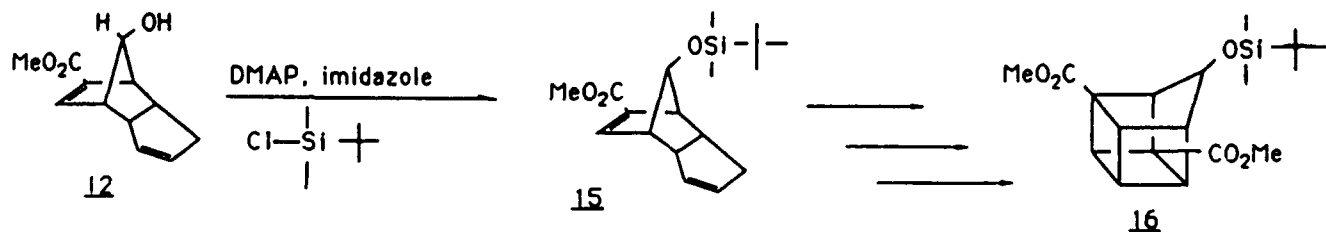
REFERENCES

1. Osawa, E.; Aigami, K.; Inamoto, Y. J. Org. Chem., 1977, 43, 2621.

2. von Schleyer, P.R.; Williams, J.E.; Blanchard, K.R.; J. Am. Chem. Soc., 1970, 92, 2377-2386.

3. Marchand, A.P.; Suri, S.C. J. Am. Chem. Soc., 1984, 49, 2041-2043.

4. Suri, S.C.; Rodgers, S.L.; Lauderdale, W.J. Tetrahedron Lett., 1988, 4031-4034.

5. Smith, W.B.; Marchand, A.P.; Suri, S.C.; Jin, P.-W. J. Org. Chem., 1986, 3052-3054.

6. Conversation with S.C. Suri, Astronautics Lab, Edwards AFB.

7. Klunder, A.J.H.; Lange, J.H.M.; Zwanenburg, B. Tetrahedron Lett.,1987, 3027-3030.

8. Cookson, R.C.; Hudec, J.; Williams, R.O. Tettrahedron Lett., 1960, 29-32.

USAF-UES RESEARCH INITIATION PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.

FINAL REPORT

SYNTHESIS OF POLY(IMIDE SILOXANE)
COPOLYMERS AND GRAFT COPOLYMERS

Prepared by: Mark A. Kirms

Academic Rank: Assistant Professor

Department and University: Chemistry Department, Southern Oregon State College

Date: December 17, 1990

Contract No.: F49620-88-0053/SB5881-0378

The following report details the progress made during the 1990 calendar year on the space polymer project started at Edwards Air Force Base in the summer of 1989. At the end of that time, poly (imide siloxane) 6 had been prepared starting from *m*-bromoaniline as shown in Scheme I. This procedure was arrived at in a serendipitous fashion after discovering that silanol 4 underwent self-condensation to give siloxane 7.

Scheme I.

Siloxane _7_ was the focus for much of the effort performed during 1990. This siloxane was viewed as the pivotal molecule in the synthesis of a wide variety of poly(imide siloxane) materials. Using McGrath's method for extending siloxane chains[1] appeared to be a straight-forward approach for initially obtaining these desired materials. There was some difficulty in preparing the required D_4 catalyst need for the "siloxation" procedure (i.e. extension of one siloxane chain into a longer siloxane chain). The catalyst prepared from the method employing toluene as a solvent seemed to work the best in terms of preparing a catalyst which was not too viscous to work with.

Several attempts were made to carry out the siloxation reaction using D_4 and the prepared catalyst under varying reaction conditions, times, and quantities of staring materials. In almost every case, no new material was observed by TLC. At this juncture it was reasoned that increasing the siloxane chain in _7_ should occur by simply heating _7_ with an appropriate amount of D_4 and a catalytic amount of powdered KOH. This procedure (Scheme II) afforded two new siloxanes whose structures appeared to be that of the trisiloxane _8_ and pentasiloxane _9_ according to their respective NMR spectra:

Scheme II.
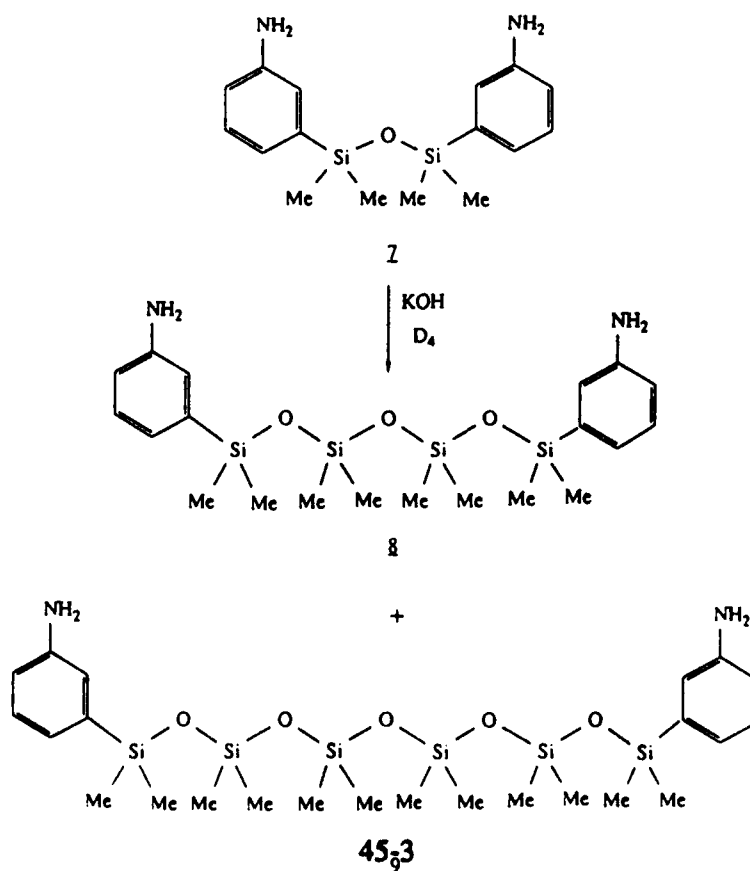
Unfortunately, these new siloxanes were obtained in very low yield ($< 5\%$) after repeated chromatography (TLC employing 1:1 EtOAc/Hexane provides both 8 and 9 from bands directly above that band corresponding to 7). In addition, a large portion (approx. 50%) of starting 7 was left unreacted. It was found that prolonged reaction time did not seem to have an effect upon the reaction (the TLC of the reaction mixture did not change significantly from 2-20 h of reaction time) nor increasing the temperature of the reaction from 120-180 °C. However, changing the proportion of KOH did have an effect, albeit detrimental, on the course of the reaction. When the proportion of KOH climbed above 15% of the starting siloxane 7, the reaction gave an intractable mixture of siloxane material with all of the starting siloxane 7 being consumed. Carrying out the reaction in dry toluene as a solvent and azeotropically removing the water showed promise of increasing the yield of 8, however, other developments concerning the preparation of 7 caused these siloxation reactions to be put on hold temporarily.

The synthesis of siloxane 7 involves initial protection of the amine as the cyclic disilazane 2. This protecting group allows for the ensuing lithium/halogen exchange reaction to occur presumably without the interference of the amine hydrogens which would otherwise be present without the of the disilazane. The protecting group was chosen because it is suppose to be impervious to alkyllithium reagents which are employed to carry out the lithium/halogen exchange [2]. However, the protecting group was found to be somewhat temperamental with regards to its behavior with alkyllithium reagents. In some instances when 2 was treated under the reaction conditions needed to converted it to silane 3, only the deprotected form of silane 3 was obtained. It was never clear whether the protecting group came off during the reaction itself or during the work-up.

The reagent needed to carry out the protection (1,1,4,4-tetramethyl-1,4-bis(N,N-dimethylaminodisilet lene) of m-bromoaniline is a costly chemical, and the question arose as to whether or not the entire protection of m-bromoaniline could be circumvented, or, at the very least, a less expensive protecting group could be found to carry out the desired transformations. For this reason, the reaction sequence shown in Scheme III was performed:

Scheme III.
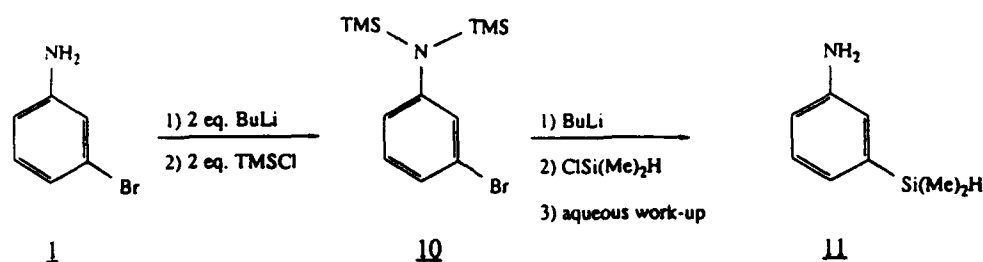
There were no intentions of isolating the disilazane 10, for it was already known that such a disilazane was less stable than the corresponding cyclic disilazane 2². Compound 10 was viewed merely as an alternative to 2 which would serve as an inexpensive way of procuring silane 11. It was already known that silane 11 could just as easily afford siloxane 7 as could the corresponding protected form of this molecule (2). Curiously, none of the desired silane was isolated from this reaction. The main product (apart from recovered starting material) isolated was the TMS-protected m-trimethylsilylaniline 12 shown below:
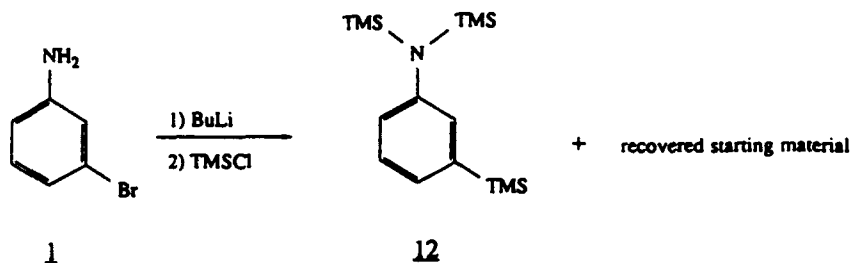
The fact that the product isolated possessed a TMS group on the aniline ring and not a Si(CH₃)₂H group seemed to imply that the ring anion is generated prior to the removal of the N-H protons by BuLi. The above result seemed to open the possibility of preparing silane 11 directly from m-bromoaniline, thus eliminating a step in the overall synthesis as well as the need for the expensive 1,1,4,4-tetramethyl-1,4-bis(N,N-dimethylamino)-disilethylene protecting group. To this end, m-bromoaniline was treated with BuLi followed by the addition of dimethylchlorosilane to give the desired silane 11 along with recovered starting material and some m-butylaniline as shown in Scheme IV. The reaction is presumed to proceed through the trianion as shown in Scheme IV. Furthermore, the assumption was made that silation would occur at carbon before occurring at nitrogen. The presence of m-butylaniline in the reaction mixture is a happenstance of bromobutane being

produced during the course of the reaction and that this alkylating reagent competes with the silating reagent for the anion. This problem could undoubtedly be circumvented by employing methyllithium to generate the trianion. In this case bromomethane would be produced as a by-product which should evaporate out of the reaction mixture.

Scheme IV.
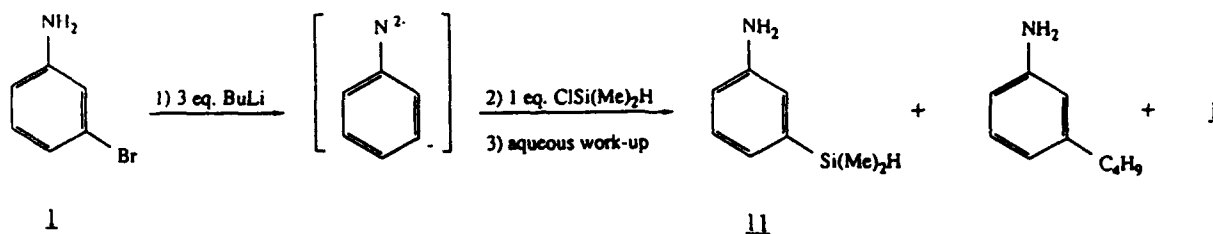
Having successfully rid the original synthesis of siloxane 7 of the need for protection of the aniline nitrogen, it became apparent that the prospect existed for preparing siloxane 7 directly from m-bromoaniline. After generating the trianion shown in Scheme IV, it seemed reasonable to assume that one could quench the trianion with 1,3-dichloro-tetramethylsiloxane instead of dimethylchlorosilane and obtain siloxane 7 directly. Scheme V details the envisioned reaction employing MeLi in place of BuLi to avoid the formation of m-butylaniline:

Scheme V.
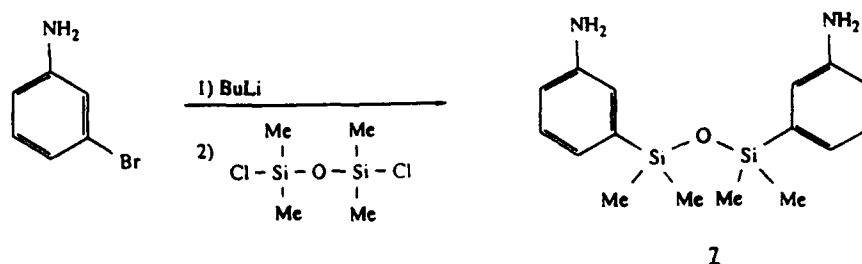
At the present time, this reaction has been preformed only once. There does appear to be a small amount of siloxane 7 which is produced as indicated by the TLC of the reaction mixture. However, this particular reagent does not perform in the same, straight-forward

manner that chlorodimethylsilane does. The quenching of the envisioned trianion is much more sluggish with the chlorosiloxane reagent than with the chlorosilane. This may perhaps be due to the inductive effect of the oxygen bonded to silicon. It is felt that this reaction can be accomplished in time, and will provide an inexpensive, one-step method for preparing siloxane 7 which is the focal point for all the ensuing poly(imide siloxane) copolymers to be prepared and tested.

References

1) a) Arnold, C. A.; Summers, J. D.; Chen, Y. P.; Bott, R. H.; Chen, D.; McGrath, J. E.; *Polymer*, **1989**, *3*, 986.

   b) Arnold, C. A.; Chen, D.; Chen, Y. P.; Graybeal, J. D.; Bott, R. H.; Yoon, T.; McGrath, J. E.; *Polym. Mater., Sci. Eng.*, **1988**, *60*.

2) Gugenheim, T. L. *Tetrahedron Letters*, **1984**, *25*, 1253.

MINI-GRANT RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.

FINAL REPORT

Prepared by: Faysal A. Kolkailah, Ph.D., P.E.

Academic Rank: Professor

Department and Aeronautical Engineering

 University: Cal Poly, San Luis Obispo

Research Location: Aeronautical Engineering
        California Polytechnic State University
        San Luis Obispo, CA 93407

USAF Researcher: Jim A. Koury

Date: 7 November, 1990

Contract No: S-210-10MG-127

FINITE ELEMENT ANALYSIS FOR COMPOSITE STRUCTURES

by

Faysal A. Kolkailah, Ph.D., P.E.

ABSTRACT

The scope of this investigation was to develop a good working finite element model using **CAEDS** to calculate stress distribution and displacements to support three independently separate projects. **CAEDS** was employed to do finite element analysis of multiple material multilayered cylindrical pressure vessel, compression of advance composite materials, and tension of notched advance composite material. The nodal displacement and element stress concentration were visually and analytically investigated using **CAEDS**. **CAEDS GRAPHICS** provided the visual capabilities and FES gave the calculated results in tabular form which allowed thorough investigation and comparison with the results of each project. The results from **CAEDS** were below the anticipated accuracy for the scope of the project.

ACKNOWLEDGEMENTS

I would like to thank the Edwards Air Force Astronautic Laboratory for sponsorship of this research. Also, I wish to express my thanks to Universal Energy Systems for this directional and administrative help and support.

I would like to take this opportunity to express my deep gratitude to both Mr. Jim Wanchek and Mr. Jim Koury for their technical support and help during the course of this research effort. I enjoyed very much working with such a fine group in an excellent environment of teamwork.

INTRODUCTION

The use of finite element analysis has become widespread in industry. This is a result of the high demand for reduced cost and low cycle time in design. The need for a good finite element method that will calculate the stress distribution and displacements of a structure or a specimen calls for the use of CAEDS.

CAEDS is a Computer Aided Engineering Design System developed by Structural Design Research Corporation (SDRC) and distributed by IBM as a product called CAEDS. It is also found under the I-deas by General Electric. CAEDS is a design system that combines capabilities to model an object, create a ꞈite element mesh on the object, translate the information to a finite element solver, and display the results on a CADAM scope. There are six major modules within CAEDS and a user can lease those which will be needed. CAEDS can also interact with other engineering tools such as CADAM, CATIA, and NASTRAN. The ease of using one major tool to combine the capabilities of several engineering tools greatly improves efficiency. CAEDS can drastically reduce the cost of designing parts and design cycle time.

For the scope of this project, three of the six modules of CAEDS were used. The GRAPHICS module offered interactive menu-driven finite element modeling using the pre-processor programs Model Creation and Enhanced Mesh Generation. GRAPHICS interfaces with CADAM for transfer of model drawing. The post-

processing was done with Output Display which allows the user to display the Finite Element Analysis (FEA) results, and gives the capability to interactively define the output to be displayed. Output Display showed the user node shapes, stresses, strains, strain energy, temperatures, reaction forces and kinetic energy. These results can be displayed using several display forms: deformed geometry, contours, XY plots, report, and arrow plots.

The second module FINITE ELEMENT SOLVER (FES) was interfaced with CAEDS GRAPHICS. FES is a general purpose linear, 3-D, finite element analysis program which addresses static, dynamic, and heat transfer problems. FES has pre-processing and post-processing capabilities and can easily interface with CAEDS GRAPHICS. Although FES does not have a large spectrum of applications it does have the advantage of being easy to use.

CAEDS offers many convenient UTILITIES to increase design speed and efficiency. The File Translator was used to write the model information such as node and element data in a format readable by the FES. This utility greatly reduces the time spent preparing a FES job. The Data Loader allowed the output from FES to be read into a universal file which CAEDS GRAPHICS can understand. Data Loader has the ability to allow the user to write their own interface with any finite element code. The system allows the user to view the results on a CADAM scope.

For the scope of this investigation, CAEDS was utilized to provide finite element results for three independently separate projects. The first project used a CAEDS model to solve for the stress distribution and displacements during

compression of composite specimen. For this project, the **CAEDS** model was used to investigate the effects of compression on a test specimen. The second project looked at the stress concentration and displacement as a result of notches on the composite specimen under tension. The results from **CAEDS** were compared to the results of another finite element analysis program. The third project solved for the contribution of shrink fit stresses on the stress distribution an displacement of a multiple material multilayered cylindrical pressure vessel. The resu..s from **CAEDS** were used to compare to the results of theoretical analysis.

BACKGROUND AND OBJECTIVE

This is a follow-up to a Summer Faculty Research Program the principal investigator carried out at the Astronautics Laboratory at Edwards AFB in Summer, 1989. The main task of the principal investigator's SFRP was to conduct a survey of the finite element codes at the AL for the analysis and design of composite structures as well as to study and evaluate the existing codes at the AL.

The ultimate goal was to identify areas needed to support and enhance the analysis and the design of composite structures using the finite element techniques. As a continuation of the 1989 research, this study, "Finite Element Analysis for Composite Structures," will employ CAEDS to analyze three different applications in composite mechanics.

In the first application, the CAEDS model will be employed to analyze for stress distribution and displacements during compression of a composite specimen to investigate the effects of compression on a test specimen. The second application will consider the stress concentration and displacement due to notches in a composite specimen under tension. In the third application, the contribution of shrink fit stresses towards the stress distribution and displacement of a multilayered composite cylindrical pressure vessel is to be considered.

PROCEDURE

The finite element was created in the GRAPHICS module of CAEDS. Using Model Creation, the geometry was defined, mesh generated, and nodes and elements were attached to the mesh. The loads, restraints, material properties, and physical properties were also defined in Model Creation. FES was used to solve the model. Before the model can be solved, a file in **SUPERB** (the finite element solver used by **CAEDS**) format must be created by the File Translator. This new file was then merged with a file that contains the temporary analysis run parameters. This combined file was solved by **SUPERB** which created the universal file; this file contained the output data reformatted for Output Display GRAPHICS program. Using Data Loader the results from **SUPERB** were read into the model and the results were reviewed using the **CADAM** scopes. This procedure was the same for the three different models with the difference only in geometry and type of elements used. For the compression of advance composites and the multilayered pressure vessel, solid elements were used rather than the plane strain elements used for the tension of the notched composite.

Compression of Advance Composite Material

For this model, solid elements were used to solve for the stress distribution and displacements. The bottom plane was restrained in all translational and rotational

46-8

directions. The top surface was loaded with nodal forces in the downward direction (figure 1). This is the best set of boundary conditions that will simulate the compression of composite materials. The results from CAEDS were used to investigate the compressive properties of the composite.

Tension of Notched Composite Material

The assumption of plane strain was incorporated for this model. Using plane strain elements, two sets of boundary conditions were created. One set assumed that one end of the specimen was fixed with no translation or rotation allowed. The specimen was loaded with uniform nodal forces on the other end (figure 14). The other set assumed that the specimen was restrained at one end with the addition of a constraint on the central node which did not allow any translation in the transverse direction. The loading used for this set was the same (figure 9). The stress concentration and deformation around the notch and the deformation was observed for both boundary cases.

Multiple Material Multilayered Cylindrical Pressure Vessel

The assumptions in this model must be carefully considered to accurately simulate the conditions of the shrink fit stress inside a multilayered pressure vessel. Using solid elements, the restraints were attached to three of the nodes around the 270 degree direction on both the bottom and top surfaces of the cylinder. The restraints allowed no translation or rotation. The load case of this model was element

pressures. These pressures were dictated by the shrink fit stresses located at the surfaces of each layer (figure 22). For the scope of this project, the stress distribution from CAEDS was observed and compared with the theoretical results given in appendix B.

DISCUSSION OF RESULTS

Although **CAEDS** was used only to solve for the stress distribution and displacements, the results were used and analyzed differently for each different model. The scope of each project demanded a different point of view and analysis to understand the results of **CAEDS** more thoroughly.

Compression of Advance Composite Material

The scope of the problem in this project is to observe the compressive properties of the composite test specimen (see appendix B). Using the visual capabilities of **CAEDS**, the deformed geometry was investigated closely. As a result the modulus of elasticity and Poisson's ratio of the material, a "compressive necking" was observed to occur during loading (figure 2). After careful analysis of the stress distribution, it was quite apparent that there was a tension force at the top surface of the model (figure 3). This tension was the result of the way the material deformed. Although this tension force was not a significant factor in the fiber direction, it does play a major role in the matrix and lamina direction. Since the weakest direction of the composite material was in the lamina direction, this tension force was great enough to cause the specimen to fail.

Without the visual capabilities of **CAEDS**, the significant factor that the tension force contributes to the failure of the specimen might have been overlooked.

For this project, the numerical results of the stresses were not as important as the visual representation of the stress distribution and displacement. The visible results showed that a more careful investigation of the tension force was necessary to see if it was significant enough to be considered.

Tension of Notched Composite Materials

The scope for this project was to develop a finite element model using **CAEDS** and compared it to the finite element results used in the original project (see appendix A). This again is more of an observation and comparison of visual results. For this project the same model was solved using two different boundary conditions. The results of the displacements were quite different for each boundary condition (figures 8 & 13) but the stress distribution was almost equal (figures 5, 6, 7, 10, 11, 12). The primary goal of the project was to check where the stress concentration occurred and compare it to the given results of the original project. After careful investigation of the stress concentration and stress distribution the **CAEDS** results for both boundary conditions were very similar to the results of the original project.

Multiple Material Multilayered Cylindrical Pressure Vessel

For this project, **CAEDS** was used as a design tool. The original project provided the theoretical values of stress. These theoretical values of stresses come from the calculations done using theoretical equations (see appendix B) and were compared to the results of **CAEDS**.

Using the "XY plot" display form, a graph of the stresses was plotted versus distance between nodes (figure 15). The stresses were also visually represented to find where the stress concentration occurred (figures 16, 17, 18, 19, 20). The displacements (figure 21) were not representative of the actual displacements but were significant in showing the direction of deformation.

The small difference between CAEDS and theory was the result of the assumptions taken in creating the model. The boundary conditions played a major role in this difference. The restrained section of the CAEDS model did not translate or rotate which is very irrepresentative of the real model. Minimizing the number of restrained nodes lessen the difference in the results but some boundary conditions have to be incorporated for the model to be solve by FES successfully.

CONCLUSION AND RECOMMENDATION

As a result of using **CAEDS**, the analysis of each project was more effective. The versatility of **CAEDS** allowed the investigation of the three different projects. The different requirements of each project has shown the different possible uses of **CAEDS**. Because of its capability of being able to interact with other engineering tools, such as **CADAM**, it made **CAEDS** very easy to use and apply. As a result of investigations done for the three projects, it was seen that the results of **CAEDS** were very accurate and helpful for more efficient engineering.

REFERENCES

Barsoum, R. S., "Triangular quarter-point elements as elastic and perfectly-plastic crack tip elements," IJNME, 11(1977), pp. 85-98.

Barton, D. C. and Soden, P. D., "Short term in-plane stiffness and strength properties of CSM reinforced polyester laminates," Composites, 13, No. 1 (January 1982), pp. 66-78.

Bathe, K-J., Wilson, E. L., and Peterson, F. E., "A Structural Analysis Program for Static and Dynamic Response of Linear Systems," University of California, Berkeley, 1973.

Brebbia, C. A., Boundary Element Method for Engineers, Newnes Butterworths, 1978.

Gwaltney, R. C., Corum, J. M., Bolt, S. E., Bryson, J. W., "Experiment Stress Analysis of Cylinder-to-Cylinder Shell Models and Comparison with Theoretical Predictions," ASME paper No. 76-PVP-10, 1976.

Hinton, E. and Owen, D. R. J., Finite Element Programming, Academic Press, 1977.

Huddleston, R. L. and Dewey, B. R., "Optimization of Elastic Multilayer Cylindrical Vessels Loaded by Pressure and Radial Thermal Gradient," ASME paper No. 72-WA/PT-2, Nov. 30, 1972.

Owen, D. R. J. and Hinton, E., Finite Elements in Plasticity, Theory and Practice, Pineridge Press Limited, Swansea, U.K., 1980.

Sabbaghian, M. and Henriquez, E. J., "Optimum Design and Creep Relaxation of

    Dissimilar Material Multilayer Shrink Fitted Vessels," ASME paper No. 81-

    PVP-30, 1981.

Sabbaghian, M. and Nandan, D., "New Concepts on the Design of Multilayer

    Cylindrical Vessels Technology, Part I, Design and Analysis," 1969,

    pp. 649-657.

Shih, C. F., DeLorenzi, H. G. and German, M. D., "Crack extension modelling with

    singular quadratic isoparametric elements," Int J Fracture, 12 (1976),

    pp. 647-651.

Wang, S. S., Yau, J. F. and Corten, H. T., "A mixed-mode crack analysis of

    rectilinear anisotropic solids using conservation laws of elasticity," Int J

    Fracture, 16, No. 3 (June 1980), pp. 247-259.

Zhang, Shuangyin, "A study on interlaminar shear strength of composites,"

    Proceedings of the 4th International Conference on Mechanical Behaviour of

    Materials, Stockholm, Sweden, August 1983, 1, pp. 565-571.

Zienkiewicz, O. C., "The Finite Element Method," McGraw-Hill, 1977.

FIGURE CAPTIONS

Figure 1   Model for the compression of specimen

Figure 2   Displacement of the compression specimen

Figure 3   Stress distribution of the compression specimen

Figure 4   Mesh used for notched specimen

Figure 5   X - stress for the second boundary condition

Figure 6   Y - stress for the second boundary condition

Figure 7   Maximum shear for the second boundary condition

Figure 8   Displacement for the second boundary condition

Figure 9   Load case for the second boundary condition

Figure 10   Maximum shear for the first boundary condition

Figure 11   Y - stress for the first boundary condition

Figure 12   X - stress for the first boundary condition

Figure 13   Displacement for the first boundary condition

Figure 14   Load case for the first boundary condition

Figure 15   XY plot of stresses for the pressure vessel

Figure 16   Von Mises stress for pressure vessel

Figure 17   Maximum shear for pressure vessel

Figure 18   Z - stress for the pressure vessel

Figure 19   Y - stress for the pressure vessel

Figure 1 - Model for the compression of specimen

Figure 2 - Displacement of the compression specimen

Figure 3 - Stress distribution of the compression specimen

Figure 4 - Mesh used for notched specimen

Figure 5 - X - stress for the second boundary condition

Figure 6 - Y - stress for the second boundary condition

Figure 7 - Maximum shear for the second boundary condition

Figure 8 - Displacement for the second boundary condition

Figure 9 - Load case for the second boundary condition

Figure 10 - Maximum shear for the first boundary condition

Figure 11 - Y - stress for the first boundary condition

Figure 12 - X - stress for the first boundary condition

Figure 13 - Displacement for the first boundary condition

Figure 14 - Load case for the first boundary condition

Figure 15 - XY plot of stresses for the pressure vessel

Figure 16 - Von Mises stress for pressure vessel

Figure 17 - Maximum shear for pressure vessel

Figure 18 - Z - stress for the pressure vessel

Figure 19 - Y - stress for the pressure vessel

Figure 20 - X - stress for the pressure vessel
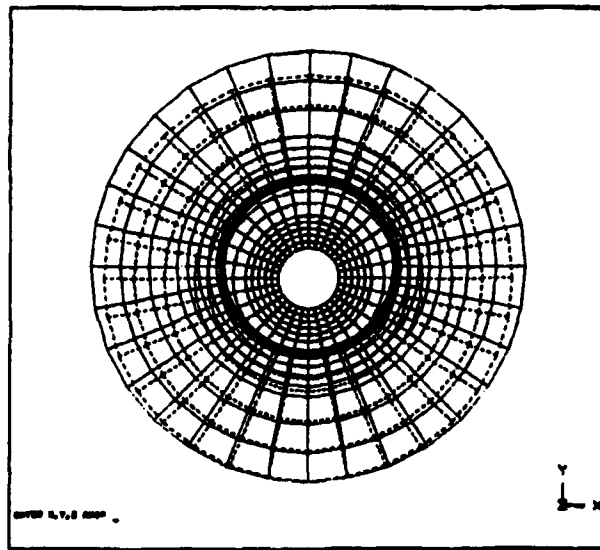
Figure 21 - Displacements for the pressure vessel
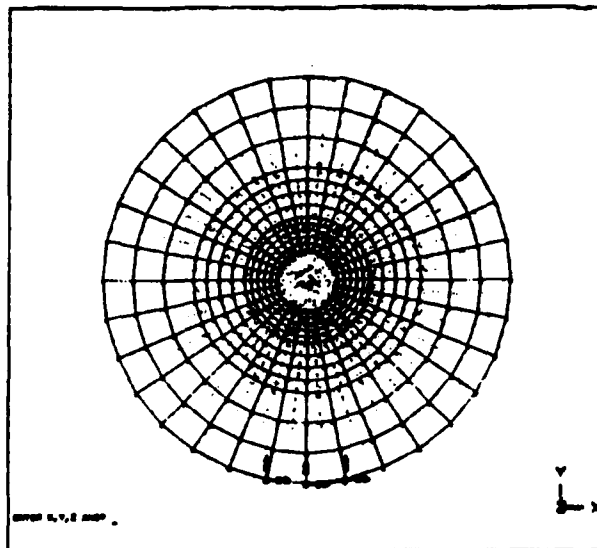
Figure 22 - Load case for the pressure vessel

APPENDICES

APPENDIX A

Tension of Notched Composite Material

Interlaminar shear fracture of chopped strand mat glass fibre-reinforced polyester laminates

S.Y. ZHANG *, P.D. SODEN † and P.M. SODEN †

(*Chinese Academy of Sciences, People's Republic of China/† UMIST, UK)

The interlaminar shear fracture of chopped strand mat glass fibre-reinforced polyester laminates has been studied both experimentally and analytically. Lap shear (double-grooved) specimens were used to measure the interlaminar shear strength and the cracking mechanism was studied using photomicrography. The finite element method was used to calculate the stress distribution along the shear surface and the mixed-mode stress intensity factors K_I and K_{II}. The length of the shear surface was found to have a significant effect on the results. Based on the experimental and analytical results, the validity of the British Standard for GRP pressure vessels (BS4994, 1973) was evaluated and the critical stress intensity factors K_{Ic} and K_{IIc} for this material were estimated.

Key words: composite materials; shear testing; interlaminar failure; stress analysis; finite element analysis; glass fibres; polyester resins

Chopped strand mat (CSM) glass fibre-reinforced polyester (GRP) is widely used in pressure vessel and pipe line systems for the chemical industry. In burst tests of pressure vessels, bending tests on pipe bends and failure of attachments to GRP vessels, interlaminar failure is often observed. Besides through-thickness tensile failure, interlaminar shear failure deserves consideration as a possible cause of delamination. In the British Standard design code for GRP pressure vessels (BS4994, 1973),[1] a test method for measuring lap shear strength of laminates is specified, but the validity of this method needs to be proved. Chiao and co-workers[2] have used this method to measure the interlaminar shear strength of Kevlar fibre laminates. The data published exhibit wide scatter. Markham and Dawson[3] proposed a simple analytical model for calculating the shear stress distribution along the shear surface. Zhang[4] has studied this problem both experimentally and analytically. The main objective of the present paper is to examine the validity of this method, particularly for CSM glass fibre-reinforced polyester laminates.

The paper consists of two parts, experimental work and a finite element analysis. In the experimental investigation, 28 single lap shear specimens of two different lengths were tested to failure under tension-shear or compression-shear loading and the crack propagation mechanism was studied using photography and photomicrography. In the finite element analysis, the stress distribution along the shear surface was calculated for different lengths of shear surface using orthotropic and isotropic material models, and crack initiation and propagation was studied.

EXPERIMENTAL DETAILS

Specimens

The test materials were flat GRP laminates produced by a commercial fabricator, Plastics Design and Engineering Limited, using Crystic 491 PA polyester resin (from Scott Bader Limited) reinforced with nine layers of 450 g m^{-2} powder-bound glass fibre CSM[5] (Fabmat PB from Fiberglass Limited). The glass fibre content was approximately 30% by weight. Twenty eight specimens in five batches were cut from three different panels (see Table 1). The shape of the specimens is illustrated in Fig. 1. The grooves were cut using a milling machine and had a width of 1.5 mm and a depth of half of the laminate thickness or a little deeper, as specified by BS4994.

Procedure

The tests were carried out in an Instron testing machine under strain-rate control. The cross-head speed was 1–2.54 mm min^{-1}. Standard Instron wedge grips were used for applying tensile loading. For compression-shear tests, the ends of the specimen were gripped in end clamps similar to those specified in BS4994 for compression testing. The clamps prevented

Table 1. Specimen details

Test method	Panel number	Number of specimens	Length of shear surface, l (mm)	Results. table number
Tension	1	5	12.1	2
Tension	2	6	13.6	3
Tension	1	7	24.4	4
Tension	3	5	12.6	5
Compression	3	5	12.1	6
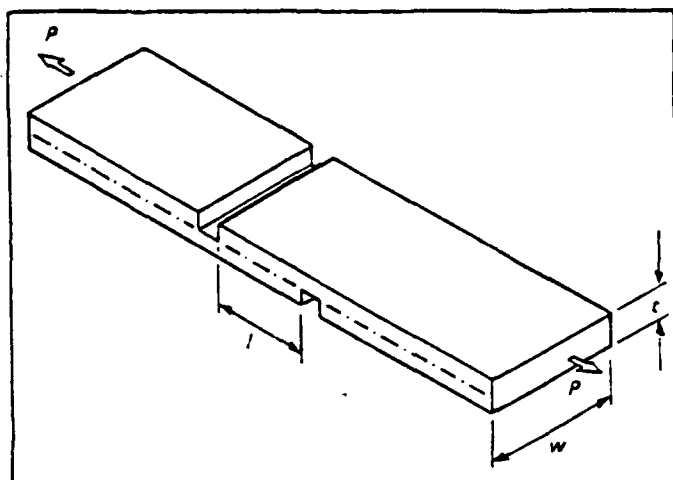
Fig. 1   Form of the notched shear specimens

Fig. 3   Load vs displacement for a compression/shear specimen (specimen no 6.1, cross-head speed 1 mm min^{-1})

Fig. 2   Load vs extension for a tension/shear specimen (specimen no 5.2, cross-head speed 1 mm min^{-1})

just before fracture to leave the specimens unbroken. Sections were cut from one of the unbroken specimens for micrographic examination. Burn-off tests[4] were conducted on samples cut from selected specimens to determine their glass content.

EXPERIMENTAL OBSERVATIONS AND RESULTS

The typical load vs cross-head displacement curve for tension–shear is shown in Fig. 2 and for compression–shear in Fig. 3. It can be seen that the curves are non-linear. In tension–shear the stiffness of the specimens decreased gradually with increasing load (see Fig. 2) until the specimen broke catastrophically. The non-linearity of these curves could be explained by the formation and propagation of cracks in the material. The crack length at various stages of loading can be measured from the photographs, see Fig. 4, or video recordings.

In both the tension–shear and compression–shear tests, a very large bending deformation was observed (see Fig 5) and large transverse normal stresses occurred at the corners of the two notches. In tension, cracks were seen to initiate at the corners at very low load levels.

the ends of the specimen rotating. Load vs elongation (cross-head displacement) curves were recorded automatically. During the tests, photographs and video recordings were taken of some specimens to record the crack propagation process and some tests were stopped
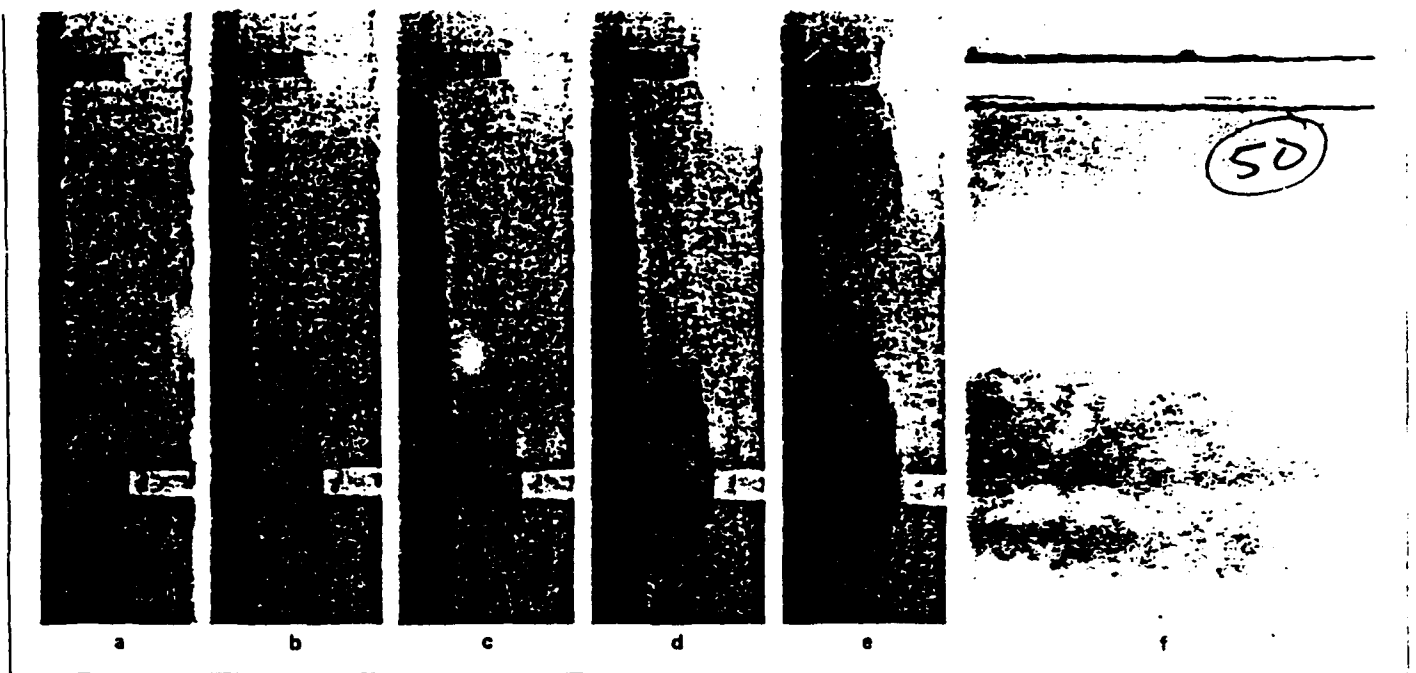
Fig. 4 (a)–(e) Crack developing in a shear specimen with increasing tensile load. The specimen is illuminated from the right. (f) as (e) but viewed from the left
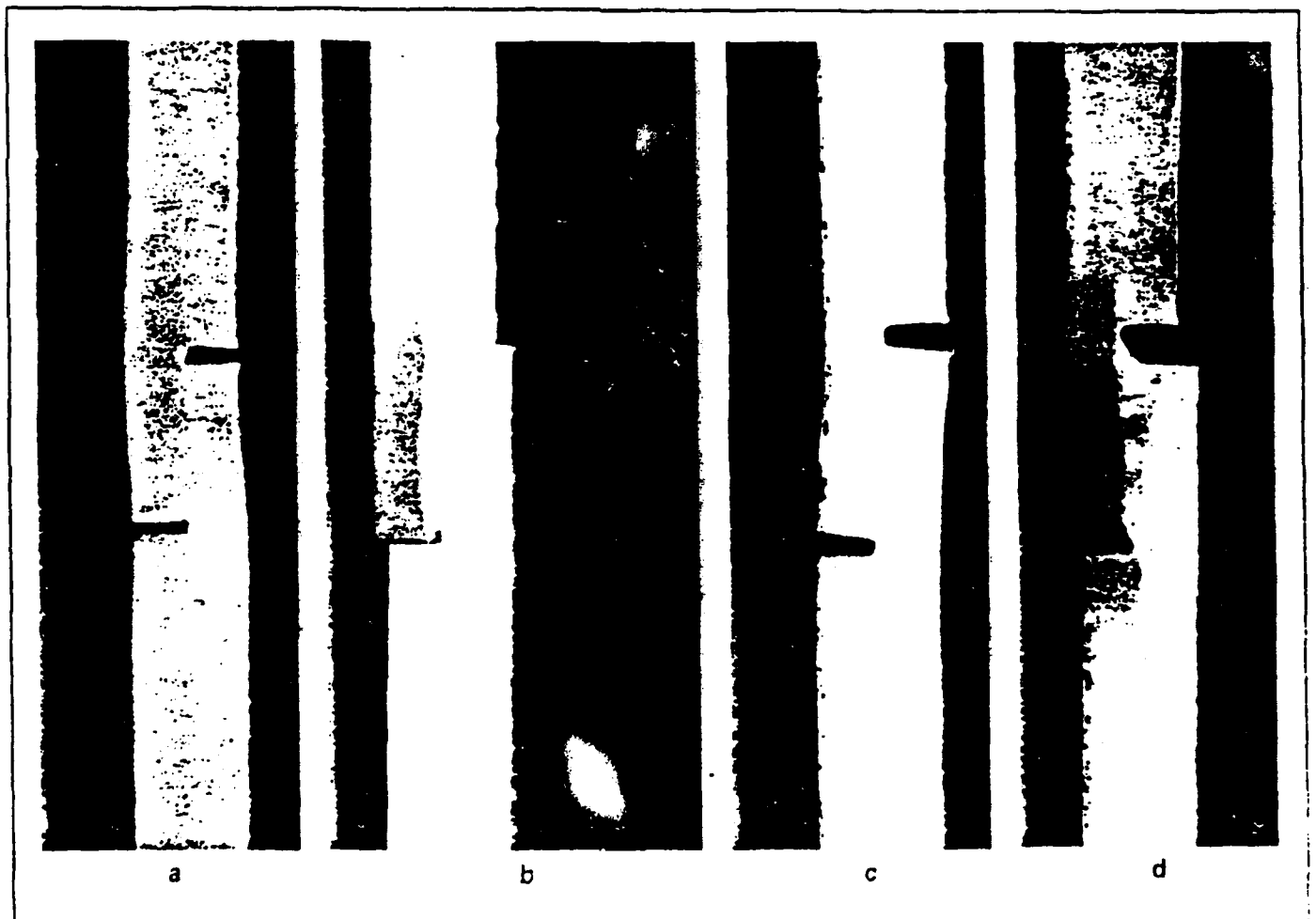
Fig. 5 (a) and (b) Shear specimen under compressive load (see Fig. 3); (c) and (d) another specimen under tensile load (see Fig. 2)

Specimen number	Thickness, t (mm)	Width, w (mm)	Length of shear surface, l (mm)	Maximum load, P (kN)	Shear strength, $\tau_{av} = P/(wl)$ (MPa)
2.1	8.4	24.95	12.0	4.67	15.6
2.2	8.3	25.0	12.1	4.81	15.9
2.3	8.25	24.9	12.5	4.63	14.8
2.4	8.4	24.9	12.1	4.45	14.75*
2.5	8.45	25.0	12.0	4.27	14.2
				Mean	15.13
				Standard deviation	0.77

Glass fraction 34.6% by weight
*Specimen unbroken — test stopped at 4.45 kN; result not included in calculation of mean

Table 3. Experimental results for tension tests on specimens cut from panel no 2

Specimen number	Thickness, t (mm)	Width, w (mm)	Length of shear surface, l (mm)	Maximum load, P (kN)	Shear strength, $\tau_{av} = P/(wl)$ (MPa)
3.1	11.68	24.95	12.5	4.21	13.49
3.2	12.15	25.08	12.8	3.56	11.08
3.3	11.79	24.92	13.5	4.21	12.5
3.4	12.2	24.95	15.0	3.64	9.71*
3.5	11.79	25.03	14.5	3.78	10.41
3.6	12.17	25.02	13.0	3.69	11.34
				Mean	11.76
				Standard deviation	1.2

Glass fraction 25.8% by weight
*Specimen unbroken — test stopped at 3.64 kN; result not included in calculation of mean

Table 4. Experimental results for tension tests on specimens cut from panel no 1

Specimen number	Thickness, t (mm)	Width, w (mm)	Length of shear surface, l (mm)	Maximum load, P (kN)	Shear strength, $\tau_{av} = P/(wl)$ (MPa)
4.1	8.69	25.8	25.0	4.14	6.42*
4.2	8.89	24.7	24.4	4.0	6.64
4.3	8.66	25.06	24.4	4.3	7.04
4.4	8.92	24.71	24.7	4.18	6.85
4.5	8.33	25.02	25.0	4.29	6.85
4.6	8.28	25.08	23.5	4.59	7.79
4.7	8.27	24.55	24.0	4.29	7.28
				Mean	7.07
				Standard deviation	0.41

Glass fraction 34.4% by weight
*Specimen unbroken — test stopped at 4.14 kN; result not included in calculation of mean

emitting faint noises and propagating in directions at a small angle to the load direction. When the cracks reached a certain (critical) length, they propagated rapidly parallel to the shear surface until reaching the other notch, then the specimen finally fractured.

The slope of the curve of compression load vs cross-head displacement (see Fig 3) decreased only at high loads and the maximum displacement of the cross-head was much less than that for the tension-shear case. The bending deformation tends to close the crack so during the stable cracking stage the crack was not as clear as in the tension-shear case. Another feature of the compression-shear tests was that as the load increased and the crack extended, puffs of smoke-like dust (see Fig 5(b)) were emitted together with faint noises

Table 5. Experimental results ...

Specimen number	Thickness, t (mm)	Width, w (mm)	Length of shear surface. l (mm)	Maximum load. P (kN)	Shear strength, $\tau_{av} = P/(wl)$ (MPa)
5.1	9.1	24.88	12.0	3.73	12.45
5.2	8.87	24.95	12.5	3.78	12.06
5.3	8.98	24.75	13.2	3.66	11.17
5.4	8.88	25.06	13.0	3.87	11.86
5.5	8.31	25.03	12.5	4.56	14.6
				Mean	12.43
				Standard deviation	1.2

Glass fraction 32.3% by weight

Table 6. Experimental results for compression tests on specimens cut from panel no 3

Specimen number	Thickness, t (mm)	Width, w (mm)	Length of shear surface. l (mm)	Maximum load, P (kN)	Shear strength, $\tau_{av} = P/(wl)$ (MPa)
6.1	8.27	25.02	11.5	5.59	19.4
6.2	8.16	24.75	13.0	5.2	16.17*
6.3	8.59	25.01	12.0	5.4	17.43
6.4	8.81	24.94	12.2	5.35	17.54
6.5	8.98	25.04	11.7	5.44	18.52
				Mean	18.22
				Standard deviation	0.93

Glass fraction 31.4% by weight
*Specimen unbroken — test stopped at 5.2 kN: result not included in calculation of mean

The test results for the 28 specimens are listed in Tables 2–6. The consistency of the tests was good and the standard deviation of the data for each group of specimens was small. The shear strength was defined[1] as the maximum load divided by the nominal area in shear ($\tau_{av} = P/wl$). For tension–shear tests the failure loads were all roughly the same so that specimens having the longer shear surface gave lower shear strengths than the shorter specimens. Note that the 12 mm long specimens (Table 2) were cut from the same panel (panel 1) as the 25 mm long specimens of Table 4. Comparing the results from Tables 2, 3 and 5 shows that panels with lower mass fractions of fibres tended to have lower shear strengths. Comparing the results of tension–shear and compression–shear tests for specimens having the same length of shear surface and cut from the same panel (panel no 3). compression loading resulted in larger shear strength than tension loading (see Tables 5 and 6).

In Fig. 4, six photographs show the crack propagation process for a tension–shear specimen with the longer shear surface. The crack initiated from the inner corner of a notch and extended initially in a direction roughly parallel to the loading direction. turned through a small angle and then propagated along the specimen again. As a result of this process a 'stair-like' crack pattern was formed. The gross angle of the crack to the load direction was approximately 10–15°. After the specimen broke. a very rough crack resulted with a large number of fibres crossing it.

Four samples were cut from a cracked but unbroken specimen and studied by photomicrography. A series of photographs were taken. From the photographs (see Figs 6 and 7) it is clear that the cracks tend to pass along the specimen through fibre-congested zones and cut at an angle across resin-rich zones. It has been shown previously[7,8] that fractures in E-glass fibre/ polyester composites tend to occur at or near to fibre/ resin interfaces. Since the fibre bundles in the chopped strand mat tend to be flattened in the plane of the laminate (see Fig. 6). the cracks propagate through these bundles in a direction roughly parallel to the surface of the laminate whereas in the resin. which is isotropic. the cracks tend to propagate in directions

Fig. 6 Micrograph showing cracks developing from the corners of a notch (the notch is at top left)

46-36
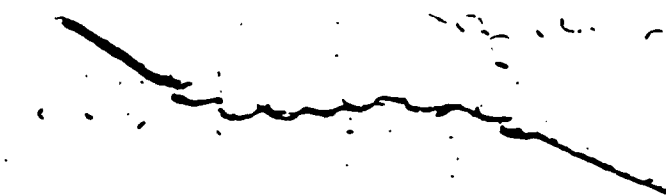
Fig. 7 Micrograph showing typical crack propagation across resin-rich areas and through a bundle of fibres

Fig. 8 Undeformed and deformed mesh pattern 1

Fig. 9 Undeformed and deformed mesh pattern 2 with transverse constraint of centre node at right-hand end

perpendicular to the directions of maximum tensile stress. This is the mechanism of the formation of stair-like crack paths. The photographs reveal clearly the inhomogeneity of the distribution of the glass fibres in the matrix resin which, together with some other imperfections such as voids, have a great influence on the path of the crack.

FINITE ELEMENT ANALYSIS

Finite element formulation

Two-dimensional eight-node quadrilateral iso-parametric elements were adopted. The formulation of this element was given in Reference 9. Collapsed triangular quarter-point singular elements[10] were used in the calculation of mixed-mode stress intensity factors K_I and K_{II}, while the normal elements were used in the analysis of stress distributions.

Material model

Two different material models were assumed: isotropic and orthotropic. In both cases the material was treated as homogeneous. For the isotropic model[11] $E = 7.0$ GPa and $v = 0.34$; whereas for the orthotropic model, $E_x = 9.81$ GPa, $E_y = 5.20$ GPa, $v_{xy} = 0.34$ and $G_{xy} = 1.43$ GPa (for 37% by weight E-glass CSM in polyester).[11,12]

Boundary conditions

One end of the specimen was assumed to be fixed with no translation or rotation allowed. The other end was subject to uniform traction. No rotation constraint was imposed on this end but in some cases the central node point was constrained not to translate in the transverse direction (see Figs 8 and 9).

Stress distribution

Stress distributions were calculated using two mesh patterns (see Figs 8 and 9). As expected, there was little difference between the results except that slightly higher stresses were obtained at the notch for mesh pattern 1 (Fig. 8) which has smaller elements at the corner of the notch. If the elements were made infinitesimally small, the stresses at a sharp corner are expected to approach infinity. Changing the constraints from those shown in Fig. 8 to those in Fig. 9 did not alter the results significantly.

Fig. 10 shows the shear stress (τ) distribution along the shear surface calculated using the mesh pattern and
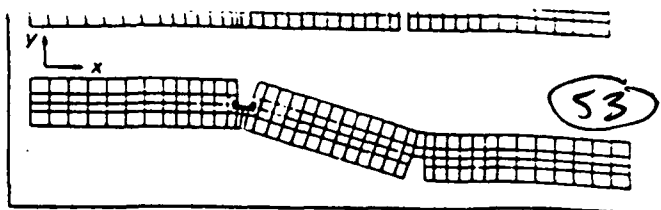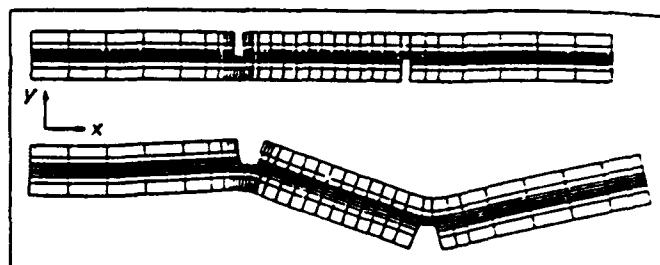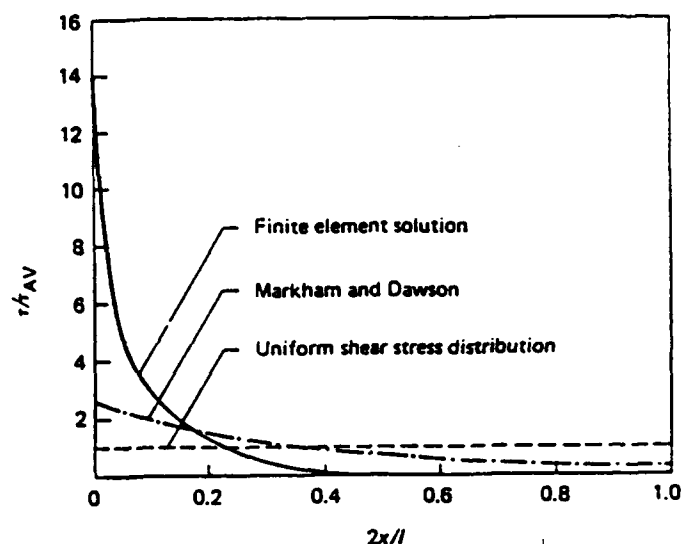
Fig. 10 Variation of shear stress with distance along the gauge length measured from the corner of the notch
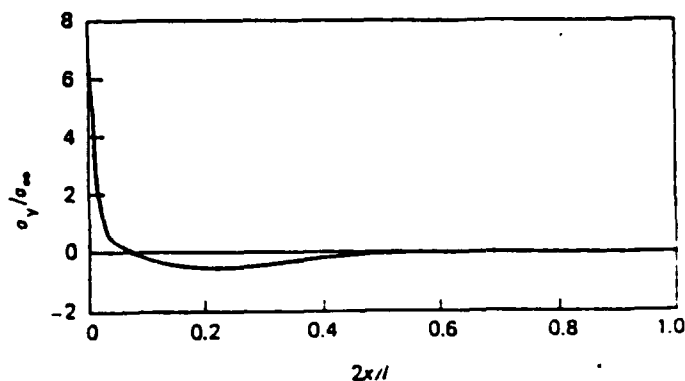
Fig. 11 Variation of transverse normal stress with distance along the gauge length

constraints of Fig. 8 and isotropic material properties. Also shown in Fig. 10 are the shear stress curves calculated using the formula given by Markham and Dawson[3] and the nominal uniform shear stress

46-37
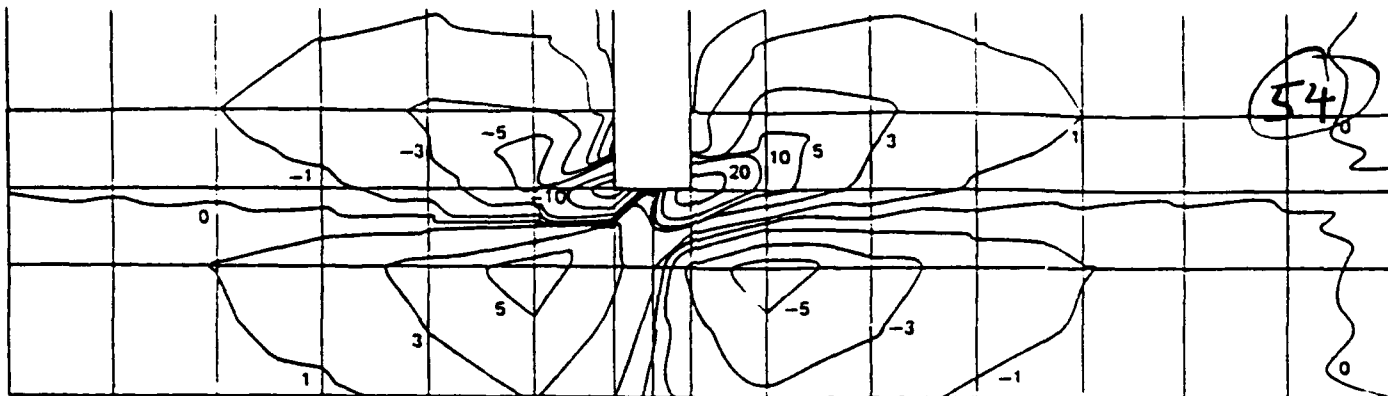
a
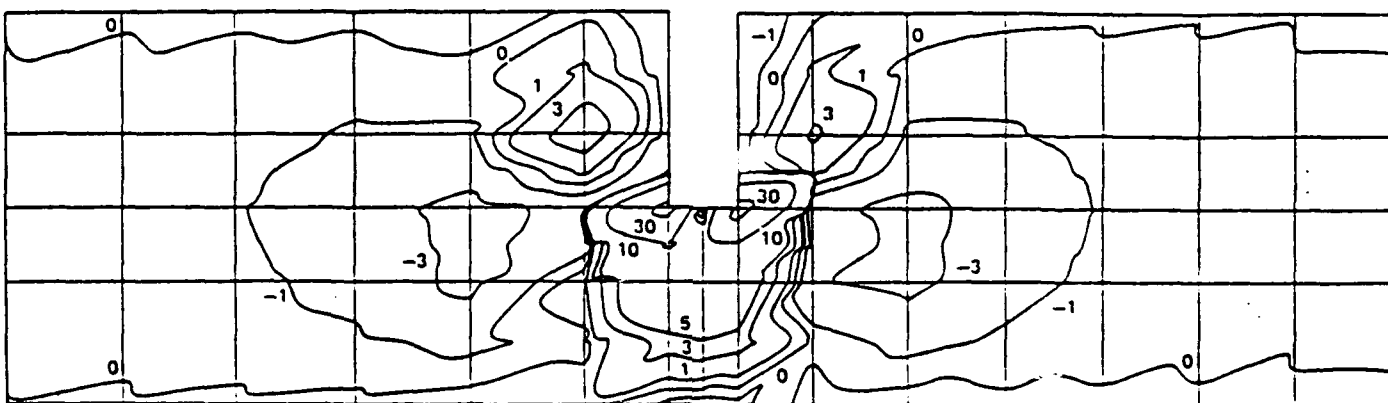
b

Fig. 12 (a) Shear; and (b) transverse normal stress distributions. Contours are lines of constant stress and numbers on the contours indicate the stress magnitude (M'a) when σ_∞ = 8 MPa. Half of the gauge length and the second notch (to the right of each figure) has been omitted

distribution as used in the calculations for shear strength. The transverse normal stress (σ_y) distribution along the shear surface is shown in Fig. 11. Both shear stress and transverse stress graphs show a very large stress concentration at the notch. It is these stress concentrations which are the main cause of the damage, cracking and failure of the specimen. Stresses calculated using the orthotropic material properties were very similar to those for the isotropic case, except that the peak shear stresses and transverse tensile stresses at the corner of the notch were 25% lower for the orthotropic material. Shear stress (τ) contours and transverse normal stress (σ_y) contours were calculated using finite element mesh pattern 1 (Fig. 8) with orthotropic material properties. The results are shown in Fig. 12.

Figs 13(a)–(c) show the theoretical shear stress distributions for three models having different shear lengths between notches; in all other respects the model specimens are the same. Mesh pattern 1 was employed with orthotropic properties. The 25 mm shear length (Fig. 13(a)) corresponds to the longer length used in the experiments and to the theoretical results discussed previously. The 12.5 mm length (Fig. 13(b)) corresponds to the shorter length used in the experiments and is also the standard length recommended in BS4994. Comparison of Figs 13(a) and (b) shows that the stress distribution near the

notch is the same in both cases, which explains why the shear strengths obtained from the experiments were larger for the shorter specimens. The method for calculating shear strength implied a uniform stress distribution. The shear distribution for the 12.5 mm length is still far from uniform, the shear stress being nearly zero over a significant fraction of the shear length. The results for the shortest (6.25 mm) length model (Fig. 13(c)) show that the shear length would have to be much shorter than the standard length for a more uniform stress distribution to be achieved. The shear stress distribution also varies with specimen thickness, thicker specimens having more uniform stress distribution.

Prediction of crack initiation

Calculations for predicting the crack initiation and the likely extent of the microcracked zone were carried out for a model with a 25 mm shear length and orthotropic elastic properties using the mesh shown in Fig. 9. Within every element the strains and stresses at 25 (5 × 5) points (including nine Gauss sampling points and eight nodal points) can be calculated. In each load step a search was conducted to find the maximum stress and its position. If the stress is larger than a critical value, microcracks will appear and fracture will occur at that position. The following criteria were employed in the computation[1]

46-38
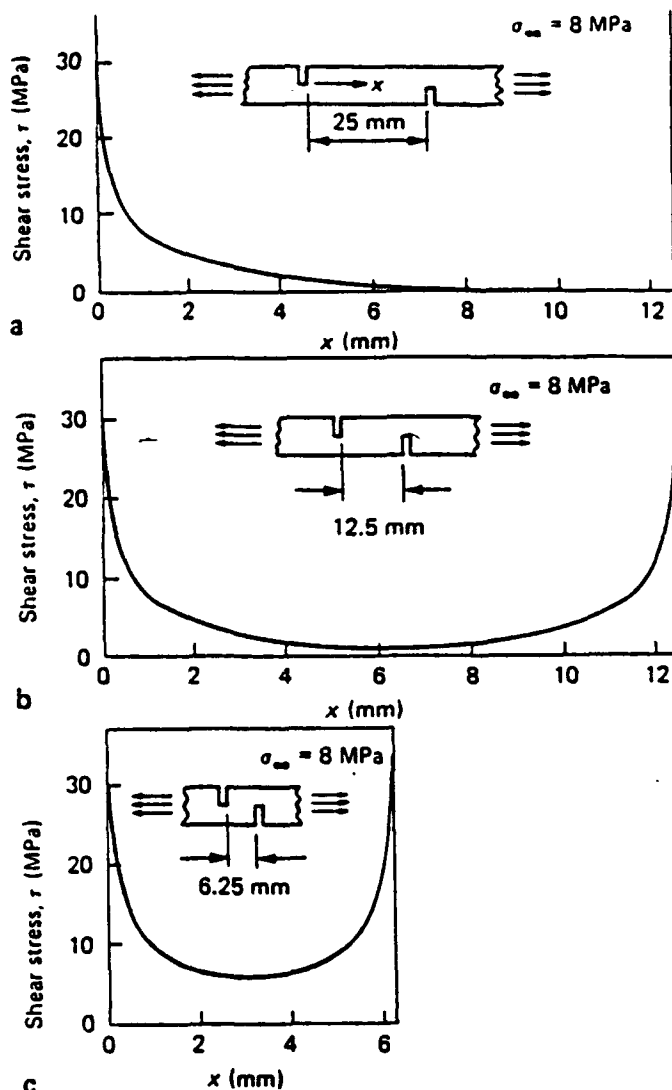
Fig. 13   Effect of shear length on shear stress distribution for shear
lengths of: (a) 25 mm; (b) 12.5 mm; and (c) 6.25 mm

1.   Maximum stress failure criteria

$$\sigma_x < \sigma_{xc}$$

$$\sigma_y < \sigma_{yc} \tag{1}$$

$$\tau < \tau_c$$

2.   Norris distortional energy failure criteria

$$\phi = [(\sigma_x/\sigma_{xc})^2 + (\sigma_y/\sigma_{yc})^2 - (\sigma_y \sigma_x/\sigma_{yc}\sigma_{xc})^2$$

$$+ (\tau/\tau_c)^2]^{1/2} < 1 \tag{2}$$

In the above two criteria σ_x, σ_y and τ are the parallel,
transverse and shear stresses respectively, and σ_{xc}, σ_{yc}
and τ_c are their critical values. From the limited
experimental data available[11-13] it was assumed that:

$$\sigma_{xc} = 120 \, \text{MPa}, \quad \sigma_{yc} = 9 \, \text{MPa} \quad \text{and} \quad \tau_c = 29 \, \text{MPa} \tag{3}$$

It must be noted that the through-thickness data is
subject to doubt.

	Criteria			
	$\sigma_x < \sigma_{xc}$	$\sigma_y < \sigma_{yc}$	$\tau < \tau_c$	$\phi < 1$
Initial failure strength,				
σ_∞ (MPa)	20.2	2.45	11.11	2.1

The initial failure point predicted by all the different
criteria was the inner corner point of the notch. Table
7 shows the initial failure strengths predicted by the
different criteria. It can be concluded from Table 7
that, as expected, the distortional strain energy
criterion, $\phi < 1$, predicts earlier damage than other
criteria (damage starting when $\sigma_\infty \simeq 2.1$ MPa in this
model), and the maximum stress criterion predicts
failure due to transverse tension at a slightly higher
strength. Although this is nominally a shear test, the
maximum stress criterion predicts failure by transverse
tension rather than shear.

When the stress (or stress factor ϕ) value meets the
failure criteria (1) or (2), the material at this point fails
and the corresponding material modulus must be
changed. For simplicity it was assumed that if $\tau \geqslant \tau_c$
then $G_{xy} = 0$; if $\sigma_y \geqslant \sigma_{yc}$ then $E_y = \nu_{yx} = 0$; if
$\sigma_x \geqslant \sigma_{xc}$ then $E_x = \nu_{xv} = 0$; and if $\phi \geqslant 1$, then
$E_x = E_y = \nu_{xy} = \nu_{yx} = G_{xy} = 0$. Other points in the
element remained unchanged. The stiffness matrix was
reformed and the whole calculation cycle repeated and
a search conducted for the next maximum stress point;
hence the crack propagation direction was found.
Using the procedure described above, the initial crack
propagation direction predicted by all the criteria (1)
and (2) is along the load direction (x), which coincides
with the observation in the micrographs, see Fig. 6.
Repeating the above calculations predicts narrow
damage zones near the corners of the notches. The
term damage zone is used rather than crack because
the analysis assumed that, at points where failure has
occurred, the material has not separated but is still
contiguous. This is an analytical rather than a physical
model. A physical representation would be a damaged
or microcracked region which exists before formation
of the macrocrack. The damaged zone predicted by the
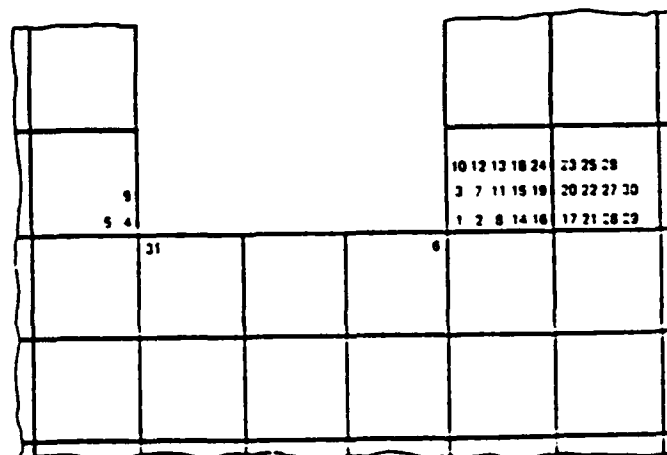
Fig. 14   Damage zone near the corner of a notch predicted using the
maximum normal stress criterion of failure (σ ... σ_y). The numbers show
position and sequence in which predicted failures occurred

46-39

Mixed-mode stress intensity factors

(56)

The mixed-mode stress intensity factors K_I and K_{II} were calculated using mesh pattern 3 shown in Fig. 16. A crack was embedded in the shear surface and 24 special elements were patched in the vicinity of the crack. Around the crack tip there were eight collapsed triangular, quarter-point singular elements. The three collapsed nodes in each element were constrained to have the same displacement. These elements possess $1/\sqrt{r}$ singularity of strain.

The crack propagation process was calculated using a constraint release technique, and the diagrams of the deformed grids at three stages of crack propagation are shown in Fig. 16(b), (c) and (d).

The relationship between K_I and K_{II} and the finite element displacement field is given as:[14]

$$U_r = K_I\sqrt{(2r/\pi)}\,Re\left[\frac{1}{m_1-m_2}\,(m_1 p_2\sqrt{\cos\theta + m_2\sin\theta}\right.$$

$$\left. -m_2 p_1\sqrt{\cos\theta + m_1\sin\theta}\right]$$

$$+ K_{II}\sqrt{(2r/\pi)}\,Re\left[\frac{1}{m_1-m_2}\,(p_2\sqrt{\cos\theta + m_2\sin\theta}\right.$$

$$\left. -p_1\sqrt{\cos\theta + m_1\sin\theta}\right] \qquad (4)$$

Fig. 15 (a) Shear; and (b) transverse normal stress distributions predicted at a line of elements next to the shear surface in a specimen with a small damage zone at the corner of the notch

maximum transverse normal stress failure criterion. $\sigma_y \leqslant \sigma_{yo}$ is shown in Fig. 14. The applied load was equivalent to a fairly low tensile load of approximately 5 MPa (see Fig. 2) and the damage zone was about 1.2 mm long. The numbers in Fig. 14 show the sequence in which failure occurred. The stress distribution of average stresses for the line of elements adjacent to the shear surface is shown in Fig. 15 for the same applied load of 5 MPa but using the failure criterion $\phi \leqslant 1$. Comparing Figs 15(a) and (b) with

Fig. 16 Mesh used to represent a specimen with a crack. (b), (c) and (d) show details of the grid representing progressive crack propagation

46-40

$$- m_2 q_1 \sqrt{\cos\theta + m_1 \sin\theta} \Big]$$

$$+ K_{II}\sqrt{(2r/\pi)} \, Re\left[\frac{1}{m_1 - m_2} \, (q_2\sqrt{\cos\theta + m_2\sin\theta} \right.$$

$$\left. - q_1\sqrt{\cos\theta + m_1\sin\theta} \right] \tag{5}$$

where m_1 and m_2 are two different roots of the following equation, and they always occur in conjugate pairs as $m_1, \overline{m}_1$ and $m_2, \overline{m}_2$:

$$a_{11}m^4 - 2a_{16}m^3(2a_{12} + a_{66})m^2 - 2a_{26}m + a_{22} = 0 \tag{6}$$

p_i and q_i are defined as:

$$p_i = a_{11}\, m_i^2 + a_{12} - a_{16}m_i \tag{7}$$

$$q_i = a_{12}m_i + a_{22}/m_i - a_{26} \tag{8}$$

where a_{ij} are compliance coefficients:

$$\epsilon_i = \sum a_{ij}\sigma_j \tag{9}$$

Using the displacement distribution on the two crack surfaces ($\theta = \pm\pi$) is the simplest way to obtain K_I and K_{II} and gives more accurate results.

$$K_I(\pi) = \sqrt{(\pi/2r_0)} \left[\frac{U_\theta(r_0\pi)}{F_I} \right] \quad (\theta = \pi) \tag{10}$$

$$K_I(-\pi) = \sqrt{(\pi/2r_0)} \left[\frac{U_\theta(r_0 - \pi)}{F_I} \right] \quad (\theta = -\pi) \tag{11}$$

$$K_I = [K_I(\pi) - K_I(-\pi)]/2 \tag{12}$$

$$K_{II}(\pi) = \sqrt{(\pi/2r_0)} \left[\frac{U_r(r_0\pi)}{F_{II}} \right] \quad (\theta = \pi) \tag{13}$$

$$K_{II}(-\pi) = \sqrt{(\pi/2r_0)} \left[\frac{U_r(r_0 - \pi)}{F_{II}} \right] \quad (\theta = -\pi) \tag{14}$$

$$K_{II} = [K_{II}(\pi) - K_{II}(-\pi)]/2 \tag{15}$$

where F_I and F_{II} can be calculated by using Equations (4)–(9).

The polar coordinate system in the vicinity of the crack tip is shown in Fig. 17.

It can be seen in Equations (10)–(15) that K_I and K_{II} depend on r_0, the distance from the crack tip to an arbitrary point on the crack surface. In the present paper, the following formulae, given in Reference 15, were used to obtain the stress intensity factors at the crack tip:

Stress intensity factor	Critical stress intensity factors
$K_I/\sigma_\infty\sqrt{a} = 1.383$	$K_{Ic} = 2.29$ MPa M$^{1/2}$
$K_{II}/\sigma_\infty\sqrt{a} = 2.285$	$K_{IIc} = 3.78$ MPa m$^{1/2}$

$$\boxed{57}$$

$$K_I = [\sqrt{\pi}/(F_I\sqrt{2})] \, [(4U_{\theta B} - U_{\theta C})/\sqrt{L}] \tag{16}$$

$$K_{II} = [\sqrt{\pi}/(F_{II}\sqrt{2})] \, [(4U_{rB} - U_{rC})/\sqrt{L}] \tag{17}$$

where $U_{\theta B}$, $U_{\theta C}$, U_{rB} and U_{rC} are circumferential and radial displacements at B and C respectively, see Fig. 18.

The values of K_I and K_{II} corresponding to the deformation shown in Fig. 16(d) are listed in Table 8, together with the critical stress intensity factors K_{Ic} and K_{IIc}, estimated for a failure load of 20 MPa and a critical crack length of 6.5 mm which are typical of the load and crack lengths observed just before failure in the experiments.

CONCLUSIONS

The single lap shear test method specified in BS4994 was examined to evaluate its suitability for measuring the interlaminar shear strength of CSM-GRP laminates. Experiments have shown that the apparent shear strength determined by this method is influenced greatly by the length between the notches, with a longer length giving lower shear strength.

During the tests, cracks started at the corners of the notches and extended to about 50% of the specimen
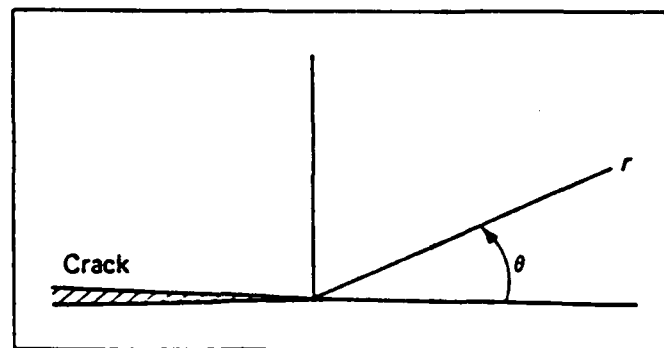
Fig. 17 Polar coordinates at the crack tip
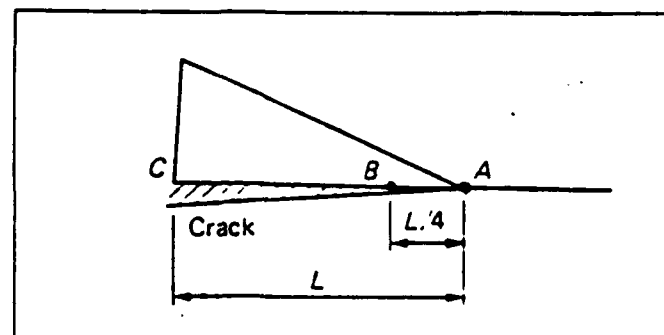
Fig. 18 Quarter-point element

46-41

micrographic investigation showed that the cracks tended to follow the fibre/matrix interface and the distribution of fibres appeared to influence the orientation of the cracks.

Large bending deformation was observed in both tension-shear and compression-shear tests and significant through-thickness normal stresses occurred. In compression-shear, these through-thickness stresses are compressive and tend to close the cracks, resulting in higher shear strengths.

The stress distribution along the shear surface obtained by finite element analysis showed large stress concentrations at the corners of the notches, which were the cause of cracking of the material and of the variation of shear strength with shear surface length.

The highly non-uniform stress distribution and the possibility of failure initiation by transverse forces makes this a poor method of measuring shear strength and the results cannot reasonably be used directly as a shear stress in engineering design calculations.

The finite element analysis of the damage zone and mixed-mode stress intensity factors gives some insight into the fracture mechanism and redistribution of stresses in the cracked specimen. Use of the maximum stress failure criterion predicted inital failure due to transverse tension rather than shear. The estimation of the critical stress intensity factors K_{Ic} and K_{IIc} may provide material properties to characterize the resistance to cracking of this material under this type of loading.

REFERENCES

1 *BS4994. Specification for Vessels and Tanks in Reinforced Plastics* (BSL London, 1973)
2 Chiao, C.C., Moore, R.L. and Chiao, T.T. 'Measurement of shear properties of fibre composites. Part 1. Evaluation of the test methods' *Composites* 9 No 3 (July 1977) pp 161-169

of fibre reinforced composites' *Composites* 6 No 5 (Sept 1975) pp 173-176
4 Zhang Shuangyin 'A study on interlaminar shear strength of composites' *Proc 4th Int Conf on Mechanical Behaviour of Materials Stockholm Sweden August 1983* 1 pp 565-571
5 *BS3496. E Glass Fibre Chopped Strand Mat for the Reinforcement of Polyester Resin Systems* (BSL London, 1973)
6 *BS2782. British Standard Methods of Testing Plastics Part 10: Glass Reinforced Plastics* (BSL London, 1977)
7 Owen, M.J. 'Fatigue testing of fibre reinforced plastics' *Composites* 1 No 6 (December 1970) pp 346-355
8 Hull, D. *An Introduction to Composite Materials* (Cambridge University Press, UK, 1981)
9 Owen, D.R.J. and Hinton, E. *Finite Elements in Plasticity. Theory and Practice* (Pineridge Press Limited, Swansea, UK, 1980)
10 Barsoum, R.S. 'Triangular quarter-point elements as elastic and perfectly-plastic crack tip elements' *IJNME* 11 (1977) pp 85-98
11 Barton, D.C. and Soden, P.D. 'Short term in-plane stiffness and strength properties of CSM reinforced polyester laminates' *Composites* 13 No 1 (January 1982) pp 66-78
12 Kitching, R., Tan, A.L. and Abu Mansour, T.M.N. 'The influence of through thickness properties on glass reinforced plastic laminated structures' *Composite Structures* 1 No 2 (1984) pp 105-151
13 Owen, M.J. 'Biaxial failure of GRP — mechanisms, modes and theories' *Proc 2nd Int Conf on Composite Structures Paisley College of Technology, Scotland 14-16 September 1983* (Applied Science Publishers, UK, 1983) pp 21-39
14 Wang, S.S., Yau, J.F. and Corten, T. 'A mixed-mode crack analysis of rectilinear anisotropic solids using conservation laws of elasticity' *Int J Fracture* 16 No 3 (June 1980) pp 247-259
15 Shih, C.F., DeLorenzi, H.G. and German, M.D. 'Crack extension modelling with singular quadratic isoparametric elements' *Int J Fracture* 12 (1976) pp 647-651

AUTHORS

S.Y. Zhang is with the Institute of Mechanics, Chinese Academy of Sciences, Beijing, People's Republic of China. P.D. Soden, to whom inquiries should be addressed, and P.M. Soden are with the Department of Mechanical Engineering, UMIST, PO Box 80, Manchester, M60 1QD, UK.

APPENDIX B

Multiple Material Multilayered Pressure Vessel

A FINITE ELEMENT PRESENTATION
OF AN OPTIMUM DESIGN FOR MULTIPLE
MATERIAL MULTILAYERED CYLINDRICAL PRESSURE VESSELS
WITH CREEP RELAXATION

F. A. Kolkailah, Professor of Aeronautical Engineering
California Polytechnic State University
San Luis Obispo, California

ABSTRACT

The subject of this paper is the application of the finite element analysis to an optimum design technique for multiple material, multilayered cylindrical pressure vessels with creep relaxation.

This optimum design technique enables the designer to calculate readily the stresses and displacements in each layer during the fabrication process or during the use of the cylinder.

SAP IV and CAEDS Programs were employed in this study. The comparisons of the finite element predictions with those obtained from Lame's solution have shown reasonably good overall agreement.

The decay in the interface pressures and the prestresses in each layer due to creep relaxation, as a function of time are obtained by employing power-function creep law. This study shows that in order to maintain the same level of maximum stress, either a gradually increasing external pressure should be applied or the vessel's inside pressure should gradually be decreased.

NOMENCLATURE

The following symbols are used in this paper:

A - radius ratio.
B - material constant controlling the rate of creep.
E - Young's modulus of elasticity.
F.S - factor of safety.
P - pressure.
r - radius at a general location.
U - displacement.
ϵ - normal strain.
ν - poisson ratio.
σ - normal stress.
τ - shear stress.
ϕ - stress ratio.
Δ - radial interference.
R - summation.

Subscripts:

c - creep.
e - elastic.
f - shrinkfit.
i - inside.
J, k or L - specific layer or specific location.
n - number of layers.
o - outside.
r - radial.
θ - tangential.
. - time derivative.

Superscripts:

m - material constant (exponent) controlling creep rate.

INTRODUCTION

One of the most common methods to build multilayered cylindrical pressure vessels is the shrink fit construction method. It uses concentric layers of relatively thick cylinders, each shrink fitted inside or outside the others, until the required design wall thickness and recompression of inner layers are reached. This technique enables the designers to get the maximum performance by combining the best properties of different materials.

A number of multilayer pressure vessels analysis based on elastic stress limits has been discussed in several investigations. Sabbaghian and Henriquez (1981), have used the maximum shear theory to optimize the design parameters. Also, in their study for the dissimilar material multilayer shrink fitted vessels, they considered the creep relaxation to present the design parameter as a function of time. Huddleston and Dewey (1972), employed the maximum distortion theory to optimize the design parameters. Sabbaghian and Nandan (1969), used the maximum shear theory to determine the optimum relations for internal pressures, radial and tangential stresses and the radial interference.

As to the finite element models, there are a large number of them in use today for both plane stress and plane strain analyses. For linear elasticity, the mathematical development of these models is simple. Since the constant strain elements have the advantages of being simple and economical, they have been selected to be employed in this study. The finite element analysis of pressure vessels has been discussed by several investigators including Zeinkiewicz (1977), Hinton and Owen (1977) and Brebbia (1978) who compared the boundary elements against finite elements and the exact elasticity solution. He reported a considerable disagreement for stresses and a reasonable agreement for the displacement. Gwaltrey, et al. (1976), in their analytical predictions for the cylinder to cylinder shell models, have employed a finite element code that used flat-plate elements which considered five degrees of freedom per node in the final assembled structures.

With time, the initial optimum stresses will change because of the creep relaxation which will reduce the original shrink fit pressure between the layers. This will change the predicted optimum stress distribution. A method to study this phenomena is needed to predict the change of the factor of safety as a function of time.

OPTIMUM DESIGN

In a cylinder subjected to inside and outside pressure, the elastic radial and tangential stresses will follow Lame's equations:

$$\begin{matrix} \sigma_\theta \\ \sigma_r \end{matrix} = \frac{P_i r_i^2}{r_o^2 - r_i^2} \left[1 - \left(\frac{r_o}{r}\right)^2 \right] \pm \frac{P_o r_o^2}{r_o^2 - r_i^2} \left[1 - \left(\frac{r_i}{r}\right)^2 \right] \quad (1)$$

Where,

- P_i, r_i — inside pressure and inside radius respectively
- P_o, r_o — outside pressure and outside radius respectively
- r — radius at a general location.

When two concentric cylinders are shrink-fitted together, the inner and outer cylinders will be subjected to external and internal pressure, respectively. The resultant stresses can be obtained from Lame's equations. In the case when a third cylinder is pressed over or into the two shell assembly, one of the shells consists of two layers, and the stresses due to shrinkage must be super imposed on those produced by previous assembly. This procedure applies at the shells of any number of layers.

Now, let us consider an optimum design of a multilayered cylindrical pressure vessel, to be made from n layers having the same elastic properties but, as most often the actual case may be, with different allowable stresses. For the jth layer with an inside radius r_j and an outside radius r_{j+1}, the inside and outside interface pressure are P_j and P_{j+1}, respectively. Let the internal pressure be P_1 and the overall pressures, P_2, P_3, etc., where P_j and P_{j+1} are functions of the internal pressure and the shrink fit pressures. The allowable tensile stress σ_{oj}, $j = 1, 2, \ldots, n$ and the shearing stress of each layer can be normalized with respect to that of first layer thus Sabbaghian (1981).

$$\phi_1 = \frac{\sigma_{o1}}{\sigma_{o1}} = 1 \text{ and } \phi_n = \frac{\sigma_{on}}{\sigma_{o1}} \quad (2)$$

To apply the maximum shear stress theory, the maximum shearing stress in each layer is equated to the allowable shearing stress for that layer.

$$r_{j\,max} = \phi_j \cdot \frac{\sigma_{o1}}{2} \qquad j = 1, 2, \ldots \quad (3)$$

From equation (3), P_1 is expressed as

$$P_{1\,max} = \frac{\sigma_{o1}}{2} \sum_{j=1}^{n} \phi_j \frac{A_j - 1}{A_j} \quad (4)$$

Where,

$$A_j = [(r_{j+1})/r_j]^2 \quad (5)$$

For the optimum value of P_{1max}, P_{1max} given by equation (4) is differentiated with respect to r_j, $j = 2, 3, \ldots n$.

$$\frac{A_1}{\phi_1} = \frac{A_2}{\phi_2} = \ldots = \frac{A_n}{\phi_n} \quad (6)$$

Then, the optimum inside pressure is given by Sabbaghian (1981)

$$P_{1opt} = \left[\frac{\sigma_{o1}}{2}\right] \left[\frac{-n}{A_1} + \sum_{j=1}^{n} \phi_j\right] \quad (7)$$

Equation (7) shows that P_{1opt} is independent of order of assembly.

For quick determination of the minimum number of layers required in a multilayer vessel for design pressure, P_1, and known material properties, equation (7) can be rearranged as

$$A_1 = [n / \sum_{j=1}^{n} \phi_j - \frac{2 P_{1opt.}}{\sigma_{o1}}] \quad (8)$$

and since A_1 is positive and has a finite value, therefore,

$$\sum_{j=1}^{n} \phi_j > \frac{2 P_{1opt.}}{\sigma_{o1}} \quad (9)$$

Equation (9) provides the minimum value of n for the optimum design. According to Sabbaghian (1981), the optimum overall pressures P_2, P_3, ... P_n at interface radii r_2, r_3, ... r_n, are given as follows:

$$P_{Lopt} \Big|_{at\ r=r_L} = \frac{\sigma_{o1}}{2} \left[- \frac{(n+1-L)}{A_1} + \sum_{j=L}^{n} \phi_j \right] \quad (10)$$

The optimum differential pressure for each layer is given by Sabbaghian (1981)

$$P_L - P_{L+1} = \frac{\sigma_{o1}}{2A_1} (\phi_L A_1 - 1) \quad L = 1, 2, \ldots n \quad (11)$$

Finally, the necessary radial interface, Δ, at each interface, could be calculated as follows:

$$\Delta_j = U_{j-1} \Big|_{r=r_j} + U_j \Big|_{r=r_j} \quad (12)$$

where at the time of assembly, U_j is given by Sabbaghian (1981)

$$U_j = \frac{r}{Ej(r_{j+1}^2 - r_j^2)} [(P_j r_j^2 + P_{j+1} r_{j+1}^2)(1 - \nu_j) + (P_j - P_{j+1}) \frac{r_{j+1}^2 r_j^2}{r^2}(1 + \nu_j)] \quad (13)$$

However, taking into consideration the fact that layers j, j-1 or both might be a part of different previously assembled subassemblies, a correction will have to be applied to equation (12) and one can write,

$$\Delta_j \Big|_{r=r_j} = \frac{\sigma_{o1} r_1 (\phi_j A_1 - 1)}{EA_1} [A_1^{j-1} \prod_{L=1}^{j-1} \phi_L]^{\frac{1}{2}} \quad (14)$$

which is independent of order of assembly.

FINITE ELEMENT ANALYSIS

The finite element program that was employed in this study is SAP IV, Sabbaghian (1981), which was chosen as being reasonably representative of currently available and widely used finite element programs. Element type four which was chosen for the finite element mesh in this study is a constant strain, two dimensional, linear elastic element with two degrees of freedom per nodal point for the final assembled structure. The optimum design analyses were employed into the finite element code to evaluate the stresses and displacements at each node. The finite element predictions were then compared with the results obtained by using Lame's solution.

At a different stage of this study the Computer Aided Engineering Design System, CAEDS, was employed as a design tool. The assumption in this model was carefully considered to accurately simulate the conditions of the shrink fit stress inside a multilayered pressure vessel. Using solid elements, the restraints were attached to three of the nodes around the 270 degree direction on both the bottom and top surfaces of the cylinder. The restraints allowed no translation or rotation. The load case of this model was element pressures. These pressures were dictated by the shrink fit stresses located at the surfaces of each layer.

CREEP RELAXATION

With time, the initially optimum design stresses changes because of the stress relaxation, and the interface pressures will decrease.

The vessel approaches the condition of a single layered vessel which will gradually decrease the safety factor for the vessel.

The stress-strain rate creep law for compound strain is given by the power function as follows

$$\dot{\epsilon}_c^* = B \, \sigma^{*m} \tag{15}$$

where, for plane strain and incompressibility conditions,

$$\epsilon^* = \text{effective strain rate} = \frac{2}{\sqrt{3}} \dot{\epsilon}_\theta = \frac{-2}{\sqrt{3}} \dot{\epsilon}_r \tag{16}$$

and

$$\sigma^* = \text{effective stress} = \frac{\sqrt{3}}{2} (\sigma_\theta - \sigma_r) = \sqrt{3} r \tag{17}$$

By substituting for the creep stress distribution in any layer j, σ_{r_j} and σ_{θ_j}, equation (17) becomes

$$\sigma_j^* = \frac{\sqrt{3}}{m_j} \frac{(r_{j+1}/r)^{2/m_j}}{(r_{j+1}/r_j)^{2/m_j} - 1} (P_j - P_{j+1}) \tag{18}$$

where m_j is the creep exponent for the material used in the j layer.

Equation (18) is valid for any r within the prescribed range. Considering equation (6) in conjunction with (18) we can write,

$$\sigma_{j max} \Big|_{r=r_j} = \frac{\sqrt{3}}{m_j} \frac{A_j^{1/m_j}}{(A_j^{1/m_j} - 1)} (P_j - P_{j+1}) \tag{19}$$

Now;

$$\epsilon_{total}^* = \epsilon_e^* + \epsilon_c^* = \text{constant} \tag{20}$$

where;

$$\epsilon_e^* = \frac{2(1 + \nu)}{3E} \sigma_c^* \tag{21}$$

By differentiating equation (21) with respect to time and substituting for $\dot{\epsilon}_e$ and $\dot{\epsilon}_c$, one can write,

$$\frac{d\sigma^*}{\sigma^{*m}} = \frac{-3EB}{2(1 + \nu)} dt \tag{22}$$

with the initial condition,

$\sigma^* = \sigma_o^*$ at $t = t_o$, we can write,

$$\sigma^{*(1-m)} = \sigma_o^{*(1-m)} + \frac{3BE(1-m)}{2(1 + \nu)} \tag{23}$$

For the j layer; equation (23) becomes

$$\sigma^{*(1-m_j)} = \sigma_o^{*(1-m_j)} + \frac{3EB_j}{2(1 + \nu)} (1-m_j) t \tag{24}$$

Combining equations (19) and (24); leads to

$$\left[\frac{\sqrt{3}}{m_j} \frac{A_j^{1/m_j}}{A_j^{(1/m_j - 1)}} (P_j - P_{j+1})_t \right]^{(1-m_j)}$$

$$= \left[\frac{\sqrt{3}}{m_j} \frac{A_j^{1/m_j}}{A_j^{(1/m_j - 1)}} (P_j - P_{j+1})_o \right]^{(1-m_j)}$$

$$- \frac{3 B_j E}{2(1+\nu)} (1-m_j) t \tag{25}$$

Substituting for the optimum differential pressure at $t = 0$, from equation (11) into equation (25), the optimum differential pressure at time t, can be written as follows:

$$(P_j - P_{j+1}) \Big|_t = (C_j - D_j t)^{\frac{1}{1-m_j}} \tag{26}$$

where

$$C_j = \left[\frac{\sigma_{ol}}{2 A_1} (\phi_j A_1 - 1) \right]^{1-m_j} \tag{27}$$

and

$$D_j = \frac{3 B_j E (1-m_j)}{2(1 + \nu)} \left[\frac{m_j (A_j^{1/m_j} - 1)}{\sqrt{3} A_j^{1/m_j}} \right]^{1-m_j} \tag{28}$$

Also, from equations (1), (2) and (26), the overall

Also, from equations (1), (2) and (26), the overall pressure differential between the layers at time t and can be found. For the nth layer, one can write,

$$P_{n+1}\Big|_t = P_1 - \sum_{j=k=1}^{n} (C_k - D_k\, t)^{1/1-mk} \qquad (29)$$

where, $P_{n+1}\Big|_t - P_o$ is the external pressure required to compensate for the stress relaxation of every layer at time t.

In practice, $P_{n+1} - P_o = 0$, therefore the allowable internal pressure at time to can be evaluated from equation (29)

$$P_1\Big|_t = P_1\Big|_t - \sum_{k=1}^{n} (C_k - D_k\, t)^{1/1-mk} \qquad (30)$$

Equation (30), shows how the internal pressure should be decreased with time in order to stay within the allowable stress σ_o.

Finally, from the factor of safety stand point, if neither the internal pressure was decreased nor the external pressure was increased with time to compensate for loss of shrink fit pressure due to creep, then the factor of safety for the vessel will decrease with time and one can write,

$$FS_t = FS_o \left[\frac{P_1\Big|_t}{P_1\Big|_{t=0}} \right]$$

or

$$FS_t = FS_o \left[\frac{\sum_{k=1}^{n} (C_k - D_k t)^{1/1-mk}}{\dfrac{\sigma_{ol}}{2} (\dfrac{-n}{A_1} + \sum_{k=1}^{n} k)} \right] \qquad (31)$$

With constant pressure operating conditions, equation (31) provides the required factor of safety for safe operation of the vessel.

EXAMPLE

Consider the design of a 20" ID pressure vessel for internal pressure of 18,000 psi. The inner layer for a typical cylinder is made of stainless steel with the following properties, $E = 24 \times 10^6$ psi, $\nu = 0.3$, $\sigma_{all} = 20,000$ psi, $m = 4.34$ and $B = 6.38 \times 10^{-27}$. As to all other n layers, they are made of carbon steel with the following properties:

$E = 24 \times 10^6$ psi, $\nu = 0.3$, $\sigma_{all} = 15,000$ psi, $m = 2.88$ and $B = 1.315 \times 10^{-17}$.

Using equation (2), $\Phi_1 = 1$ and $\Phi_2 = \Phi_3 = 0.75$

For the optimum number of layers, we can use equation (9) as follows,

$$\frac{2P_1}{\sigma_{ol}} = 1.8 \quad \text{and} \quad \sum_{j=1}^{n} \Phi_j$$

$$= 1 + 0.75 + 0.75 + \ldots$$

Now, $1.8 < 1 + 0.75 + 0.75 + \ldots$, which shows that the minimum number of layers is 3, i.e. n $\geq$ 3. Therefore, let us consider $n_{optimum}$ to be 3. From equations (8), and (6), $A_1 = 4.286$ and $A_2 = A_3 = 3.214$. From equation (5), $r_2 = 20.7$, $r_3 = 37.11$ and $r_4 = 66.53$ inches. From equation (14), the radial interferences at r_2 and r_3 are 0.0382 and 0.0685 inches respectively.

With the above radial interferences our final set of dimensions are, $r_1)_i = 10$, $r_1)_o = 20.704$, $r_2)_i = 20.70$, $r_2)_o = 37.178$, $r_3)_i = 37.11$, $r_3)_o = 66.53$ and $r_4 = 66.50$ inches.

As to the shrink-fit pressure between the shells, it was calculated as follows,

$$P_{f1} = (r_3^2 - r_2^2) \left[\frac{\Phi_2}{2} \frac{\phi_{ol}}{r_3^2} + \frac{\Phi_3}{2} (\frac{r_1}{r_4})^2 (\frac{r_4^2 - r_3^2}{r_3^2 - r_1^2}) \right.$$

$$\left. - P_1 (\frac{r_1}{r_2})^2 (\frac{r_1}{r_3^2 - r_1^2}) \right]$$

and

$$P_{f2} = (r_4^2 - r_3^2) \left[\frac{\Phi_3}{2} \frac{\sigma_{ol}}{r_4^2} - P_1 (\frac{r_1}{r_3})^2 (1/r_4^2 - r_1^2) \right]$$

Finally, for the optimum different pressure for each layer, by employing equation (11), $P_1-P_2 = 766.822$ Psi and $P_2-P_3 = 5166.822$ Psi respectively.

Now, to maintain the original optimum stress level in the first layer, one will have to decrease, with time, the internal pressure P_1 or otherwise, gradually increase the external pressure.

If the internal pressure is to remain the same, the external pressure to be added is obtained from equations (27), (28) and (29)

$$P_{n+1} - P_o \quad \begin{array}{l} = 5021 \text{ p.s.i} \\ \text{after one year} \\ t = 8760 \text{ hours} \end{array}$$

On the otherhand if the external pressure remains zero, then the internal pressure should be decreased according to equation (30)

$$P_1\Big|_{t = 8760 \text{ hours}} = P_i\Big|_{t = 8760 \text{ hours}}$$

$$= 12979 \text{ p.s.i}$$

Finally if both the internal and external pressures are maintained at original levels, then the factor of safety will decrease with time according to equation (31) and after one year (8760 hours) the factor of safety will be 72% of the origin of factor of safety.

RESULTS AND CONCLUSION

Figure 1 shows the typical assembly with the calculated dimensions. The optimum analyses were the input data to the SAP IV Code. The finite element predictions along with the calculated stresses and displacements. using Lame's equations, are presented in

Table 1. The comparison shows good agreement between the finite element results for the optimum design and the exact solution. However, the percentage error for the stresses is higher at both ends with a bigger value for the tangential stresses.

As to CAEDS analyzing, Fig. 2 shows the stresses versus distance between nodes. The small difference between CAEDS and theory is the result of the assumptions considered in creating the model. The boundary conditions played a major role in this difference. The restrained section of the CAEDS model did not translate or rotate which is very irrepresentive of the real model. Minimizing the number of restrained nodes lessen the difference in the results, but some boundary conditions have to be incorporated.

Under constant pressure conditions, the designer can predict the stress variations due to relaxation because of the loss in the shrink fit pressure, and can prescribe a program for change of pressure with time in order to maintain the original factor of safety. Alternatively, he can project the useful life of the vessel before the factor of safety becomes unacceptable.

REFERENCES

Bathe, K-J., Wilson, E.L., and Peterson, F. E., "A structural Analysis Program for Static and Dynamic Response of Linear System", University of California, Berkeley, 1973.
Brebbia, C. A., "Boundary Element Method for Engineers" Newnes Butterworths, 1978.
Gwaltney, R. C., Corum, J. M., Bolt, S. E., Bryson, J. W., "Experiment Stress Analysis of Cylinder-to-Cylinder Shell Models and Comparison with Theoretical Predictions", ASME paper No. 76-PVP-10, 1976.
Hinton, E. and Owen, D. R. J., "Finite Element Programming", Academic Press, 1977.
Huddleston, R. L. and Dewey, B. R., "Optimization of Elastic Multilayer Cylindrical Vessels Loaded by Pressure and Radial Thermal Gradient", ASME paper No. 72-WA/PT-2, Nov. 30, 1972.

Sabbaghian, M. and Henriquez, E. J., "Optimum Design and Creep Relaxation of Dissimilar Material Multilayer Shrink Fitted Vessels", ASME paper No. 81-PVP-30, 1981.
Sabbaghian, M. and Nandan, D., "New Concepts on the Design of Multilayer Cylindrical Vessels Technology, Part I, Design and Analysis", 1969, pp. 649-657.
Zienkiewicz, O. C., "The Finite Element Method", McGraw-Hill, 1977.
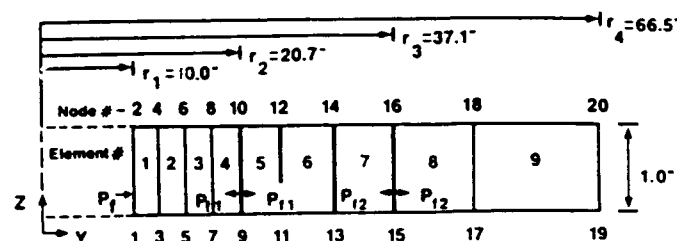
Fig. 1 Typical Assembly with the Calculated Dimensions
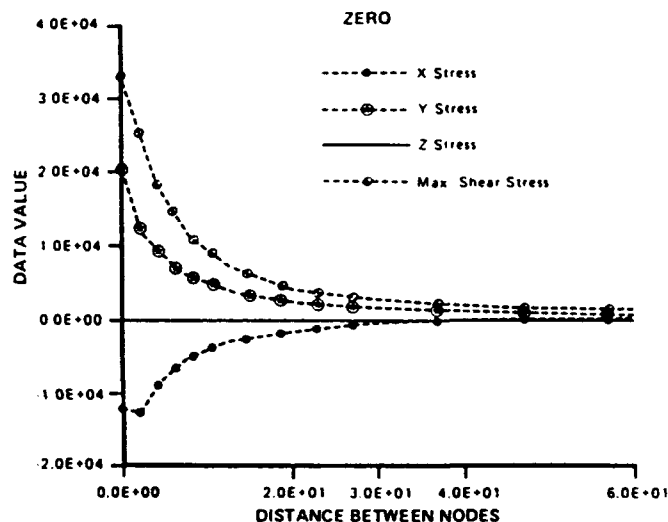
Fig. 2 CAEDS: Stresses Versus Distance Between Nodes

Table 1 Exact and F-E Solutions for Stresses and Displacements

Solutions \ radius	10"	18"	20.7"	37.1"	66.5"
Exact Solution	2,018.54	-4,901.44	-5,654.90	-1,820.90	5,295.87
σ_θ F-E Solution	2,188.20	-4,867.90	-5,636.00	-1,791.00	5,436.27
% Error	8.4	.68	0.33	1.60	2.60
Exact Solution	-18,000.00	-11,080.30	-10,326.56	-5,858.61	0.00
σ_r F-E Solution	-17,574.70	-11,042.10	-10,284.10	-5,790.60	113.70
% Error	2.30	0.34	0.40	1.60	---
Exact Solution	0.0031	-0.00118	0.000875	-0.0124	0.0146
U F-E	0.0036	-0.00110	0.0009	-0.0140	0.0101

Final Report - Research Initiation Program

ROBUST CONTROL OF LARGE FLEXIBLE STRUCTURES
USING REDUCED ORDER MODELS

sponsored by

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

conducted by

Universal Energy Systems, Inc.
Contract No. F49620-88-C-0053\SB 5881-0378

Submitted by

S. Vittal Rao, Ph.D.
Professor of Electrical Engineering
University of Missouri-Rolla
Rolla, MO 65401

ABSTRACT

The design and implementation of control strategies of large, flexible structures presents challenging problems. One of the difficulties in controller design arises from the incorrect knowledge of the stuructu al parameters. Another problem is to approximate a high order, complex mathematical model of the system with low order, simpler models. The balanced realization, optimal projection and L_2/H_∞ model reduction techniques are employed for deriving reduced order models for experimental grid structure. Numerical solution algorithms are developed for solving the necessary conditions of optimality associated with optimal projection method. The reduced order modeling methods intorduces non-minimum phase (NMP) zeros in the mathematical model of the experimenatl grid. The influence of NMP zeros on the performance of the closed loop LQG/LTR controller is investigated. A reduced order robust controller is implemented on the grid structure by using Max 100 real time computer. Experimental closed loop performance of the grid is obtained for various parameter variations.

I. INTRODUCTION

The design and implementation of control strategies of large, flexible structures presents challenging problmes. One of the difficulties in controller design arises from the incorrect knowledge of the structural parameters. In order to guarantee stability and provide satisfactory performance in the presence of model uncertainties, the linear quadratic Gaussian with loop transfer recovery (LQG/LTR) design procedure is used to design robust controlles. Another problem in the control of a large flexible structure is to approximate a high order, complex mathematical model of the system with low order simpler models. The resulting reduced order models are employed for designing robust controllers and implemented on the structure. This methodology will provide simplicity of implementation and reduction in complexity of hardware requirements.

A large number of model reduction methods are available in the literature. This problem has received considerable attention in recent years. The notable methods are eigenvalue truncation methods [1], modal and component cost analysis [2], balance-truncation method [3], optimal projection methods [4,5] and combined L_2/H_∞ model reduction methods [6]. The reduced order models are developed primarily for controller design purposes. The real test of the controller derived using any of the model reduction methods is based on its performance relative to the controller synthesized using the original system representation. The derivation of reduced order models for large, flexible structures presents problems due to the presence of a large number of closely-spaced, lightly damped coupled modes. Another difficulty in model reduction arises from 'spillover' from high-order discarded modes. The eliminated higher order modes can destablize the control system. The balanced realization, optimal projection and L_2/H_∞ model reduction methods are investigated and a suitable method is selected for determining simplified models for the structures with rigid and flexible modes. We have also evaluated error bounds in the model reduction methods.

A control strategy which can guarantee stability and provide satisfactory performance in the presence of model uncertainties is called a robust controller. These uncertainties may include modelling error between the control synthesis model and the actual system, parameter variations and the effects of various system/sensor distrubances on system performance. In

controller design for a large flexible structure, there is always a trade off between performance and robustness. When the highest performance is required, the robustness characteristics will be ciminished.

Among the various design methods for robust controllers, the linear quadratic Gaussian with loop transfer recovery (LQG/LTR) deisng procedure has many advantages. This methodology is particularly suitable for the control of large flexible structures due to the considerable modelling inaccuracy that inherently exists in the mathematical models. Sundarajan, et al [7], Joshi and Armstong [8], Yadavalli [9] have successfully applied LQG/LTR methodology for the design of robust controllers for flexible structures. A procedure for designing LQG/LTR controllers based on reduced order models of the structure, along with the spillover problem, is presented in this report.

Hyland and Bernstein [10] have presented a technique for model reduction by minimizing the steady state error between the outputs of the original system and reduced order model. A detailed description of this procedure along with numerical solution algorithm is presented in this report.

Recently Hadad and Bernstein [6] have developed a L_2/H_∞ reduction procedure for deriving reduced order models by minimizing a quadratic model reduction criterion subject to a prespecified (worst case frequency domain) H_∞ constraint on the frequency domain error between original and reduced order models. When the H_∞ constraint is sufficiently relaxed, this method will reduce to optimal projection procedure. A brief description of L_2/H_∞ model reduction method is included in this report.

The reduced order models are employed to design LQG/LTR controllers for the Astronautics Laboratory (AL) experimental grid structure. The reduced order modeling methods introduces non-minimum phase (NMP) zeros in the mathematical representation of the grid structure. The influence of NMP zeros on the performance of the closed loop LQG/LTR controller is also investigated. A reduced order robust controller is implemented on the grid structure by using Max 100 real time computer. The closed loop performance of the grid is recorded to illustrate the effects of parameter variations and sensor noise on the results.

II.   DESCRIPTION AND STRUCTURAL MODELLING OF EXPERIMENTAL GRID

An experimental grid structure is built at the Astronautics Laboratory, Air Force System Command (AL) to develop a simple ground test bed for future large flexible structures. The principal objective of this research facility has been to achieve satisfactory agreement between theoretical results and experimental measurements [11].

The two-dimensional experimental structure shown in Fig. 1 consists of a 5' x 5' grid made up of 2" wide, 1/8" thick aluminum strips. At every point where the vertical and horizontal strips cross each other, they are connected by four rivets, thus effectively removing any play at the joints. The grid hangs vertically down, being cantilevered at the top to a large I-beam anchored to a cinder block wall.

The structural vibrations are monitored using high sensitivity, low mass piezoelectric accelerometers. Permanent magnet DC motor torquers are used as actuators for the grid. The grid can be exited by an electrodynamic shaker or a quartz impulse hammer.
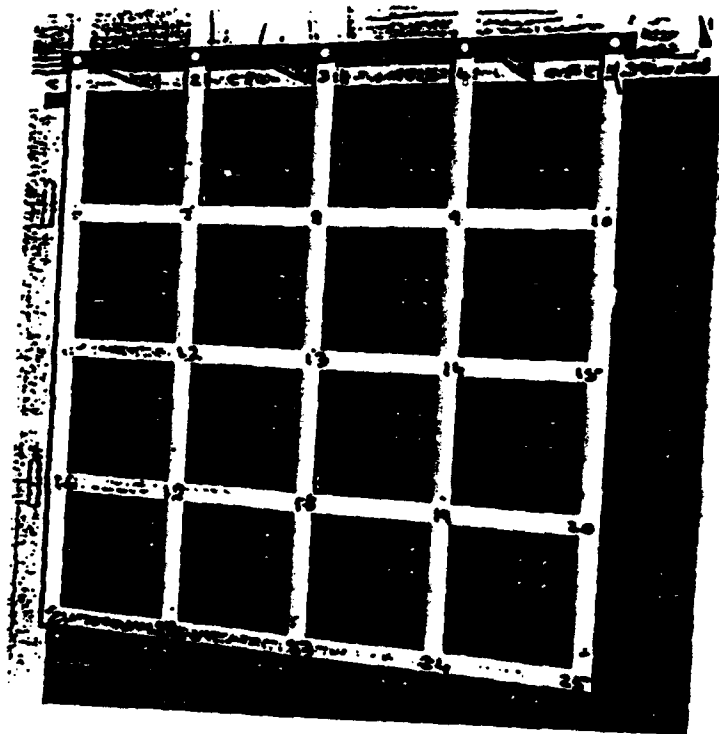
Fig. 1. A L Experimental Grid.

Structural Modelling

A model of the structure which relates inputs and outputs is needed for control design purposes. A theoretical mathematical model of the structure is developed using finite element methods. The natural frequencies of the Nastran finite element model are compared with the experimentally obtained natural frequencies of the grid.

In order to design robust controllers for the structure we need a state variable representation of the grid. A procedure for realization of a state variable model for finite element model is given below.

The experimental grid structure is represented by 75 degrees of freedom (DOF) finite element-model. Each grid node has three DOFS: translation in z direction and rotations about horizontal and vertical axes. Equation of motion of the grid are given by

$$[m]\ddot{q} + [c]\dot{q} + [k]q = F(t) \qquad \ldots(1)$$

The vector q(t) represents the physical coordinates of the structure.

A 10-mode model representation of the grid is evaluated by using a linear transformation and truncation

$$q = \Phi\eta \qquad \ldots(2)$$

where η = Modal coordinates of the structure
Φ = eigenvectors

Modal representation of the grid structure is given by

$$\ddot{\eta} + \text{diag}(2\zeta\omega_1,\ldots 2\zeta\omega_{10})\dot{\eta} + \text{diag}(\omega_1^2,\ldots \omega_{10}^2)\eta = \Phi^T Au(t) \qquad \ldots(3)$$

where F(t) = Au(t)
A = A matrix contains information about location and type (x - torquer / y - torquer) of actuators
ζ = A uniform damping of 0.0025 is assumed for every mode.

A state variable model corresponding to Eq. (3) is given by

$$\dot{x} = \begin{bmatrix} 0 & I \\ -\text{Lambda} & -\text{Zetamat} \end{bmatrix} x + \begin{bmatrix} 0 \\ \Phi^T A \end{bmatrix} u \qquad \ldots(4)$$

$$\dot{x} = Ax + Bu \qquad \ldots(5)$$

where $\quad \dot{x} = \begin{bmatrix} \eta \\ \dot{\eta} \end{bmatrix}$

$$\text{Lambda} = \text{diag}(\omega_1^2, \omega_2^2, \ldots \omega_{10}^2)$$

$$\text{Zetamat} = \text{diag}(2\zeta\omega_1, \ldots 2\zeta\omega_{10})$$

Norris et al [12] have presented a procedure for determining the dynamic effect of gravity on low frequency accelerometer measurements. These gravity effects are incorporated in the output equation as given below:

The output of the ith accelerometer is given by

$$a_i = \ddot{q}_i + g(\cos\alpha_i)q'_i \qquad\qquad , i = 1, 2, \ldots m. \qquad \ldots(6)$$

where α_i is the angle between vertical and the tangent to the structure in equilibrium at accelerometer location p_i.

From eq (3) and (4)

$$\ddot{\eta} = -\text{diag}\left(\omega_1^2, \ldots \omega_{10}^2\right)\dot{\eta} - \text{diag}(2\zeta\omega, \ldots 2\zeta u_\omega)\eta + \Phi^T A u \qquad \ldots(7)$$

$$\ddot{\eta} = \begin{bmatrix} 0 & I \end{bmatrix} \begin{bmatrix} 0 & I \\ -\text{Lambda} & -\text{Zetamat} \end{bmatrix} \begin{bmatrix} \eta \\ \dot{\eta} \end{bmatrix} + \begin{bmatrix} 0 & I \end{bmatrix} \begin{bmatrix} 0 \\ \Phi^T A \end{bmatrix} u$$

$$\ddot{\eta} = \begin{bmatrix} 0 & I \end{bmatrix}[Ax] + \begin{bmatrix} 0 & I \end{bmatrix} Bu \qquad \ldots(8)$$

From eq (2)

$$\ddot{q} = \Phi\ddot{\eta} = \Phi\begin{bmatrix} 0 & I \end{bmatrix}Ax + \Phi\begin{bmatrix} 0 & I \end{bmatrix}Bu \qquad \ldots(9)$$

consider the accelerometer reading given by

$$a_i = \ddot{q}_i + g(\cos\alpha_i)q'_i$$

the angle $\alpha_i = 0$ for grid structure

$$q'_i = \Phi\eta' = \begin{bmatrix} \Phi & 0 \end{bmatrix}x \qquad \ldots(10)$$

Hence output equations given by

$$y = SUVMAT*a \qquad \qquad ...(11)$$

where SUVMAT = standard unit vectors for accelerometers.

From equations (through (11),

$$y = SUVMAT\{\Phi[0 \quad I]Ax + \Phi[0 \quad I]Bu + [\Phi \quad 0]x\}$$

$$y = SUVMAT\{\Phi[0 \quad I]Ax + [\Phi \quad 0]x\}$$

$$+ SUVMAT\{\Phi[0 \quad I]Bu\} \qquad \qquad ...(12)$$

$$y = Cx + Du \qquad \qquad ...(13)$$

$$\text{where } C = SUVMAT\left\{\Phi[0 \quad I]A + \begin{bmatrix} \Phi & 0 \\ 0 & 0 \end{bmatrix}\right\}$$

$$D = SUVMAT\{\Phi[0 \quad I]B\}$$

The transfer function of the grid structure is given by

$$G(s) = C(sI - A)^{-1}B + D \qquad \qquad ...(14)$$

The first ten-natural frequencies of the grid determined by using NASTRAN finite element model and experimental values are given in Table 1.

TABLE 1 Comparison of Natural Frequencies

Mode #	NASTRAN (Hz)	Experiment (Hz)
1	0.7784	0.762
2	1.90	1.810
3	4.1641	4.110
4	4.9618	5.150
5	6.3587	6.220
6	.7318	10.750
7	1 .8196	11.05
8	11.3227	11.50
9	13.8646	13.80
10	16.092	16.6

III. MODEL REDUCTION TECHNIQUE USING OPTIMAL PROJECTION

Hyland and Bernstein [10] have presented a technique for model reduction by minimizing the error between the outputs of the original system and reduced order model. The necessary conditions for optimality will result in two modified Lyapunov equations which determine the structure of the reduced order model. These Lyapunov equations are coupled by a projection parameter whose rank is equal to the order of the reduced model. A detailed description of the procedure is presented in this section. The principal references for this section are Hyland and Bernstein [10] and Ngo[13].

<u>Statement of Problem</u>:

For an nth order stable, controllable and observable system

$$\dot{x} = Ax + Bu \qquad \qquad ...(15)$$
$$y = Cx \qquad \qquad ...(16)$$

where u is white noise with spectral density V>0, find a n_mth order reduced model

$$\dot{x}_m = A_m x_m + B_m u \qquad \qquad ...(17)$$
$$y = C_m x_m \qquad \qquad ...(18)$$

by minimizing the model reduction error criterion

$$J(A_m, B_m, C_m) \triangleq \lim_{t \to \infty} E\left[(y - y_m)^T R(y - y_m)\right] \qquad \qquad ...(19)$$

where R is symmetric, positive definite matrix.

A conceptual representation of original system and reduced order model is given in Figure 2.
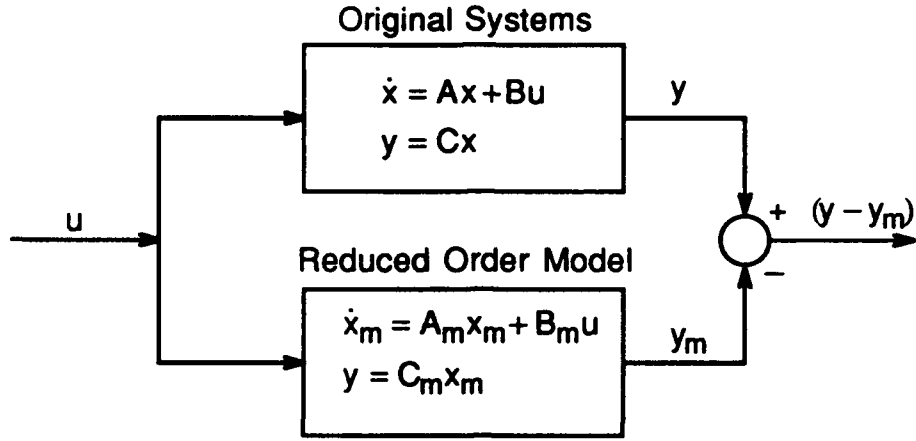
Figure 2. Representation of original and reduced order models

Consider the augmented system

$$
\begin{bmatrix} \dot{x} \\ \dot{x}_m \end{bmatrix} = \begin{bmatrix} A & 0 \\ 0 & A_m \end{bmatrix} \begin{bmatrix} x \\ x_m \end{bmatrix} + \begin{bmatrix} B \\ B_m \end{bmatrix} u
$$

$$
\dot{\tilde{x}} = \tilde{A}\ \tilde{x} + \tilde{B}u \qquad \qquad ...(20)
$$

The model reduction criterion J can be written as

$$
J = \lim_{t \to \infty} E\left[[Cx - C_m x_m]^T R [Cx - C_m x_m] \right]
$$

$$
J = \lim_{t \to \infty} E\left[\tilde{x}^T \tilde{R} \tilde{x} \right] \qquad \qquad ...(21)
$$

where
$$
\tilde{R} = \begin{bmatrix} C^T R C & -C^T R C_m \\ -C_m^T R C & C_m^T R C_m \end{bmatrix} \qquad \qquad ...(22)
$$

By using the formula $p_{1 \times n} q_{n \times 1} = tr(qp)_{n \times n}$ Equation (21) is modified to

$$
J = \lim_{t \to \infty} E\left\{ tr\tilde{x}\tilde{x}^T \tilde{R} \right\} \qquad \qquad ...(23)
$$

By interchanging the trace and expectation operation, we have

$$J = \lim_{t \to \infty} \text{tr}\left\{E\left[\tilde{x}\tilde{x}^T\right]\tilde{R}\right\} \qquad \ldots(24)$$

$$J = \lim_{t \to \infty} \text{tr}\left[\tilde{Q}(t)\tilde{R}\right] \qquad \ldots(25)$$

where $\qquad \tilde{Q}(t) = E\left[\tilde{x}\tilde{x}^T\right] \qquad \ldots(26)$

If $\tilde{A}$ is stable, then $\tilde{Q}(t)$ has limiting solution, $\lim_{t \to \infty} \tilde{Q}(t) = \tilde{Q}$ satisfying equation

$$\tilde{A}\tilde{Q} + \tilde{Q}\tilde{A}^T + \tilde{V} = 0 \qquad \ldots(27)$$

where $\qquad \tilde{V} = \tilde{B}\tilde{B}^T = \begin{bmatrix} BVB^T & BVB_m^T \\ B_m VB^T & B_m VB_m^T \end{bmatrix} \qquad \ldots(28)$

Now the problem is transformed into minimizing

$$J(A_m, B_m, C_m) = \text{tr } \tilde{Q}\tilde{R} \qquad \ldots(29)$$

subject to a constraint

$$\tilde{A}\tilde{Q} + \tilde{Q}\tilde{A}^T + \tilde{V} = 0 \qquad \ldots(30)$$

The constrained minimization problem can be solved by introducing the Lagrange multiplier $(\tilde{P})$ and modifying the performance index as

$$L = \text{tr}\left[\tilde{Q}\tilde{R} + \left(\tilde{A}\tilde{Q} + \tilde{Q}\tilde{A}^T + \tilde{V}\right)\tilde{P}\right] \qquad \ldots(31)$$

The necessary conditions for optimality are:

(i) $\qquad \dfrac{\partial L}{\partial \tilde{Q}} = \tilde{R}^T + \tilde{A}^T\tilde{P} + \tilde{P}\tilde{A} = 0 = \tilde{A}^T\tilde{P} + \tilde{P}\tilde{A} + \tilde{R} \qquad \ldots(32)$

(ii) $\qquad \dfrac{\partial L}{\partial \tilde{P}} = \tilde{A}\tilde{Q} + \tilde{Q}\tilde{A}^T + \tilde{V} = 0 \qquad \ldots(33)$

(iii) $\qquad \dfrac{\partial L}{\partial A_m} = 2\dfrac{\partial}{\partial A_m}\left(\tilde{A}\tilde{Q}\tilde{P}\right) = 0 \qquad \ldots(34)$

Let the matrices $\tilde{Q}$ and $\tilde{P}$ ar partitioned as

$$\tilde{Q} = \begin{bmatrix} Q_1 & Q_{12} \\ Q_{12}^T & Q_2 \end{bmatrix} \quad ; \quad \tilde{P} = \begin{bmatrix} P_1 & P_{12} \\ P_{12}^T & P_2 \end{bmatrix} \qquad \ldots(35)$$

Substitution of (35) in (34) and evaluation of partial derivative yields

$$\frac{\partial L}{\partial A_m} = 2\left(P_{12}^T Q_{12} + P_2 Q_2\right) = 2\left(Q_{12}^T P_{12} + Q_2 P_2\right) = 0 \qquad \ldots(36)$$

(iv) $\quad \dfrac{\partial L}{\partial B_m} = \dfrac{\partial}{\partial B_m}\left(tr\tilde{V}\tilde{P}\right) = 0 \qquad \ldots(37)$

Substitution of (28) and (35) in (37) and simplification gives

$$\frac{\partial L}{\partial B_m} = 2\left(P_{12}^T B + P_2 B_m\right)V = 0 \qquad \ldots(38)$$

Another condition of otimality is given by

(v) $\quad \dfrac{\partial L}{\partial C_m} = \dfrac{\partial}{\partial C_m}\left(tr\tilde{Q}\tilde{R}\right) = 0 \qquad \ldots(39)$

Substitution of (22) and (35) in (39) and simplification yield

$$\frac{\partial L}{\partial C_m} = 2R(C_m Q_2 - C Q_{12}) = 0 \qquad \ldots(40)$$

After expanding Equations (32) and (33) in the submatrices, the necessary conditions for optimality can be summarized as

$$A^T P_1 + P_1 A + C^T R C = 0 \qquad \ldots(41)$$

$$A^T P_{12} + P_{12} A_m - C^T R C_m = 0 \qquad \ldots(42)$$

$$A_m^T P_2 + P_2 A_m + C_m^T R C_m = 0 \qquad \ldots(43)$$

$$A Q_1 + Q_1 A^T + B V B^T = 0 \qquad \ldots(44)$$

$$A Q_{12} + Q_{12} A_m^T + B V B_m^T = 0 \qquad \ldots(45)$$

$$A_m Q_2 + Q_2 A_m^T + B_m V B_m^T = 0 \qquad \ldots(46)$$

$$Q_{12}^T P_{12} + Q_2 P_2 = 0 \qquad \ldots(47)$$

$$P_{12}^T B + P_2 B_m = 0 \qquad \ldots(48)$$

$$C_m Q_2 - C Q_{12} = 0 \qquad \ldots(49)$$

47-12

We can notice that A_m, B_m and C_m are independent of Q_1 and P_1 and thus Equations (41) and (44) can be ignored. In order to simplify the necessary conditions for optimality the following variables are defined:

$$\hat{Q} \triangleq Q_{12}Q_2^{-1}Q_{12}^T \qquad \qquad ...(50)$$

$$\hat{P} \triangleq P_{12}P_2^{-1}P_{12}^T \qquad \qquad ...(51)$$

$$G \triangleq Q_2^{-1}Q_{12}^T \qquad \qquad ...(52)$$

$$\Gamma \triangleq -P_2^{-1}P_{12}^T \quad , \quad M \triangleq Q_2P_2 \qquad \qquad ...(53)$$

$$\tau \triangleq G^T\Gamma = -Q_{12}Q_2^{-1}P_2^{-1}P_{12}^T \qquad \qquad ...(54)$$

$$\tau_\perp \triangleq I_n - \tau \qquad \qquad ...(55)$$

The necessary conditions of optimality (41-49) are simplified by using the defined variables (50-55) and the results of model reduction technique based o optimal projection method are summarized in the following theorem:

Theorem

Suppose (A_m, B_m, C_m) solves the optimal model-reduction problem. Then there exist nonnegative definite matrices $\hat{P}$, $\hat{Q}$ such that, for some factorization of $\hat{Q}\hat{P} = G^T M\Gamma$, the matrices A_m, B_m and C_m are given by

$$A_m = \Gamma A G^T \qquad \qquad ...(56)$$

$$B_m = \Gamma B \qquad \qquad ...(57)$$

$$C_m = C G^T \qquad \qquad ...(58)$$

and such that with $\tau = G^T\Gamma$, the following conditions are satisfied:

$$A\hat{Q} + \hat{Q}A^T + BVB^T - \tau_\perp BVB^T\tau_\perp^T = 0 \qquad \qquad ...(59)$$

$$A^T\hat{P} + \hat{P}A + C^T RC - \tau_\perp^T C^T RC\tau_\perp = 0 \qquad \qquad ...(60)$$

$$\text{rank}\left(\hat{Q}\right) = \text{rank}\left(\hat{P}\right) = \text{rank}\left(\hat{Q}\hat{P}\right) = n_m \qquad \qquad ...(61)$$

where $\qquad \tau_\perp = I_n - \tau$

The minimum model reduction cost is given by

$$J_{min} = tr\left[\left(Q_1 - \hat{Q}\right)C^T RC\right] \qquad ...(62)$$

$$J_{min} = tr\left[BVB^T\left(P_1 - \hat{P}\right)\right] \qquad ...(63)$$

Numerical Solution of Optimal Projection Equations for Model Reduction

In order to find reduced order models, we need to solve coupled Liapunov type equations. Essentially, two distinct approaches have been developed for solving the optimal projection equations.: (1) Iterative procedure using balancing [13] or component cost analysis [2] and (2) Homotopic continuation method [14]. In this report the iterative procedure is described.

Numerical Algorithm

Step 1: Determine the balanced realization (A_b, B_b, C_b) and Σ from the general state variable representation (A, B, C) of the original system.

Let $\qquad \Sigma = diag\left(\sigma_1, \sigma_2, ... \sigma_{n_m}\right) \qquad ...(64)$

with $\qquad \sigma_1 \geq \sigma_2 \geq ... \geq \sigma_n \qquad ...(65)$

where σ_i's are Hankel Singular values

Select the order of the reduced model, n_m such that

$$\sigma_{n_m} >> \sigma_{n_m+1} \qquad ...(66)$$

then $\qquad \Sigma_1 = diag\left(\sigma_1 ... \sigma_{n_m}\right) \qquad ...(67)$

$$\Sigma_2 = diag\left(\sigma_{n_m+1} ... \sigma_n\right) \qquad ...(68)$$

Step 2: Choose the initial values of $\hat{Q}$ and $\hat{P}$ as

$$\hat{Q}_0 = \hat{P}_0 = \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix} \qquad ...(69)$$

then

$$\tau_0 = \left(\hat{Q}_0 \hat{P}_0\right)\left(\hat{Q}_0 \hat{P}_0\right)^{\#} = \begin{bmatrix} I_{n_m} & 0 \\ 0 & 0 \end{bmatrix} \qquad ...(70)$$

Step 3: Let $\tau_{\perp i} = I_n - \tau_i$ for $i = 0, 1, \ldots$ ...(71)

Solve the following equations for $\hat{Q}_{i+1}$ and $\hat{P}_{i+1}$

$$A_b \hat{Q}_{i+1} + \hat{Q}_{i+1} A_b^T + B_b B_b^T - \tau_{\perp i} B_b B_b^T \tau_{\perp i} = 0 \qquad \ldots(72)$$

$$A_b^T \hat{P}_{i+1} + \hat{P}_{i+1} A_b + C_b^T C_b - \tau_{\perp i} C_b^T C_b \tau_{\perp i} = 0 \qquad \ldots(73)$$

for $i = 0, 1, 2, \ldots$

Step 4: For a given value of i, the solutions of (72) and (73), $\hat{Q}_{i+1}$ and $\hat{P}_{i+1}$ may have the rank greater than n_m and one of the requirements of optimal projection is that

$$\text{rank } \hat{Q} = \text{rank } \hat{P} = n_m \qquad \ldots(74)$$

To reduce the rank of $\hat{Q}_{i+1}$ and $\hat{P}_{i+1}$, find the singular value decomposition, such that

$$\hat{Q}_{i+1} = U_{i+1}^q \overline{Q}_{i+1} V_{i+1}^{q^T} \qquad \ldots(75)$$

$$\hat{P}_{i+1} = U_{i+1}^p \overline{P}_{i+1} V_{i+1}^{p^T} \qquad \ldots(76)$$

Partition $\hat{Q}_{i+1}$ and $\hat{P}_{i+1}$ as

$$\hat{Q}_{i+1} = \begin{bmatrix} \overline{Q}_{11} & 0 \\ 0 & \overline{Q}_{22} \end{bmatrix} \; ; \; \hat{P}_{i+1} = \begin{bmatrix} \overline{P}_{11} & 0 \\ 0 & \overline{P}_{22} \end{bmatrix} \qquad \ldots(77)$$

The rank of $\hat{Q}_{i+1}$ and $\hat{P}_{i+1}$ is reduced by defining

$$\overline{Q}_{i+1} = U_{i+1}^q \overline{\overline{Q}}_{i+1} V_{i+1}^{q^T} \qquad \ldots(78)$$

$$\overline{P}_{i+1} = U_{i+1}^p \overline{\overline{P}}_{i+1} V_{i+1}^{p^T} \qquad \ldots(79)$$

where $\overline{\overline{Q}}_{i+1} \triangleq \begin{bmatrix} \overline{Q}_{11} & 0 \\ 0 & 0 \end{bmatrix} \; ; \; \overline{\overline{P}}_{i+1} = \begin{bmatrix} \overline{P}_{11} & 0 \\ 0 & 0 \end{bmatrix}$ ...(80)

Step 5: Update the value of the projection parameter

$$\tau_{i+1} = \left(\overline{\hat{Q}}_{i+1} \overline{\hat{P}}_{i+1} \right) \left(\overline{\hat{Q}}_{i+1} \overline{\hat{P}}_{i+1} \right)^{\#} \qquad \ldots(81)$$

Step 6:   Define the error

$$e_1 = \frac{\left\| \overline{\hat{Q}}_{i+1} - \overline{\hat{Q}}_i \right\|}{\| \Sigma \|} \quad \text{and} \quad e_2 = \frac{\left\| \overline{\hat{P}}_{i+1} - \overline{\hat{P}}_i \right\|}{\| \Sigma \|} \qquad \ldots(82)$$

and $e = \max (e_1, e_2)$

If e is less than a prespecified value go to Step 7, otherwise go to Step 3

Step 7:   Decompose $\left(\overline{\hat{Q}} \overline{\hat{P}} \right)$ as

$$\left(\overline{\hat{Q}} \overline{\hat{P}} \right) = \Phi^{-1} \begin{bmatrix} \Lambda & 0 \\ 0 & 0 \end{bmatrix} \Phi \qquad \ldots(83)$$

Then $\tau = \Phi^{-1} \begin{bmatrix} I_{n_m} & 0 \\ 0 & 0 \end{bmatrix} \Phi \qquad \ldots(84)$

$$\Gamma = \begin{bmatrix} I_{n_m} & 0 \end{bmatrix} \Phi \quad \text{and} \quad G = \begin{bmatrix} I_{n_m} & 0 \end{bmatrix} \Phi^{-T} \qquad \ldots(85)$$

The reduced order model matrices are given by

$$A_m = \Gamma A_b G^T \qquad \ldots(86)$$
$$B_m = \Gamma B_b \qquad \ldots(87)$$
$$C_m = C_b G^T \qquad \ldots(88)$$

Ngo[14] has developed a program code in Matrixx to solve the optimal projection equations.

IV. COMBINED L_2/H_∞ MODEL REDUCTION METHOD

Haddad and Bernstein [11] have developed a procedure for deriving a reduced order model by minimizing a quadratic model reduction criterion subject to a prespecified H_∞ (worst-case frequency domain) constraint on the frequency domain error between original and reduced order models. The solution of necessary conditions for optimality for this method involves a coupled system of equations consisting of four modified Riccati equations. When the H_∞ constraint is sufficiently relaxed, these coupled equations will reduce to the optimal projection equations for model reduction given by Hyland and Bernstein [10]. A brief description of L_2/H_∞ model reduction method is included in this section. The key reference for this work is Haddad and Bernstein [11].

Problem Statement:

For an nth order stable, controllable and observable system

$$\dot{x} = Ax + BDw \qquad \qquad ...(89)$$
$$y = Cx \quad ; \qquad \qquad ...(90)$$

where w is white noise with unit spectral density.
Find a n_m ($n_m < n$) order reduced model

$$\dot{x}_m = A_m x_m + B_m Dw \qquad \qquad ...(91)$$
$$y_m = C_m \, x_m \quad ; \qquad \qquad ...(92)$$

which satisfies the following criteria:
(i)   the reduced order model is asymptotically stable.
(ii)   the L_2 (quadratic) model-reduction criterion

$$J(A_m, \, B_m, \, C_m) \; \underset{=}{\triangle} \; \lim_{t \to \infty} E\left\{ (y - y_m)^T R(y - y_m) \right\} \qquad ...(93)$$

is minimized.
(iii)   The transfer function of the reduced-order model lies within a radius - γ, H_∞ neigbourhood of the full order system,

$$\|H(s) - H_m(s)\|_\infty \le \gamma \qquad \ldots(94)$$

where
$$H(s) \triangleq E\ C(sI - A)^{-1}\ B\ D \qquad \ldots(95)$$

$$H_m(s) \triangleq E\ C_m(sI - A_m)^{-1}\ B_m D \qquad \ldots(96)$$

$$\gamma > 0 \text{ is a given constant}$$

The procedure for dertermining L_2/H_∞ reduced order model is conceptually similar to the procedure developed for fixed order robust controller.
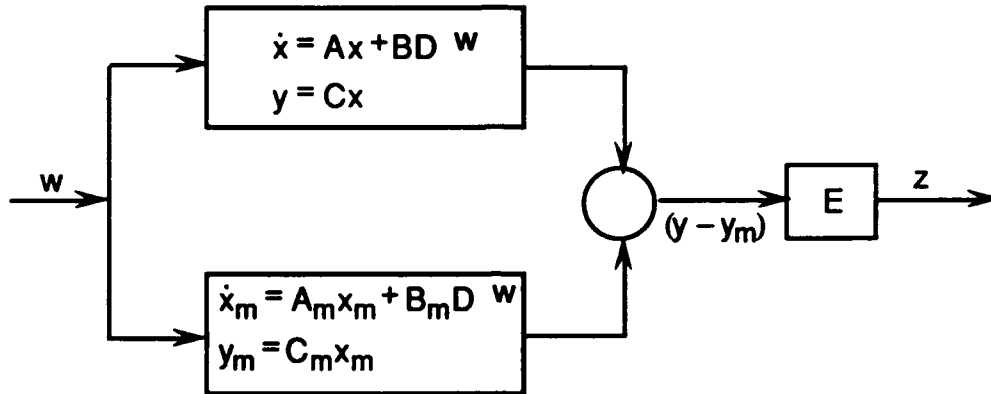
Consider the augmented system:

$$\begin{bmatrix} \dot{x} \\ \dot{x}_m \end{bmatrix} = \begin{bmatrix} A & 0 \\ 0 & A_m \end{bmatrix} \begin{bmatrix} x \\ x_m \end{bmatrix} + \begin{bmatrix} BD \\ B_m D \end{bmatrix} w$$

$$\dot{\tilde{x}} = \tilde{A}\ \tilde{x} + \tilde{D}\ w \qquad \ldots(97)$$

Let
$$z = \tilde{E}\ \tilde{x} \qquad \ldots(98)$$

where
$$\tilde{E} = E\tilde{C} = [EC \quad -EC_m] \qquad \ldots(99)$$

Then the transfer function between w and z is

$$\tilde{H}(s) = \tilde{E}\ \left(sI - \tilde{A}\right)^{-1}\ \tilde{D} \qquad \ldots(100)$$

Now consider the L_2 model reduction criterion

$$J(A_m, B_m, C_m) = \lim_{t \to \infty} \mathbf{E}\left\{(y - y_m)^T R(y - y_m)\right\} \qquad \dots(101)$$

$$= \lim_{t \to \infty} \mathbf{E}\left\{\tilde{x}^T \tilde{R} \tilde{x}\right\} \qquad \dots(102)$$

$$= \lim_{t \to \infty} \mathbf{E}\left\{\mathrm{tr}\ \tilde{x}\ \tilde{x}^T R\right\}$$

Interchanging expectation and trace operation, we have

$$J(A_m, B_m, C_m) = \lim_{t \to \infty} \mathrm{tr}\ \mathbf{E}(\tilde{x}\ \tilde{x}^T)\cdot\tilde{R}$$

$$J(A_m, B_m, C_m) = \mathrm{tr}\ \tilde{Q}(t)\tilde{R} \qquad \dots(103)$$

where
$$\tilde{Q}(t) = \mathbf{E}\left(\tilde{x}\ \tilde{x}^T\right) \qquad \dots(104)$$

$$\tilde{R} = \tilde{E}^T \tilde{E} \begin{bmatrix} C^T RC & -C^T RC_m \\ -C_m^T RC & C_m^T RC_m \end{bmatrix} ; \quad R = E^T E \qquad \dots(105)$$

and the steady state value of $\tilde{Q}(t)$ satisfies a covariance equation

$$\tilde{A}\ \tilde{Q} + \tilde{Q}\tilde{A} + \tilde{V} = 0 \qquad \dots(106)$$

where
$$\tilde{V} = \tilde{D}\ \tilde{D}^T = \begin{bmatrix} BVB^T & BVB_m^T \\ B_m VB^T & B_m VB_m^T \end{bmatrix} \qquad \dots(107)$$

$$V = D\ D^T \qquad \dots(108)$$

In order to accommodate H_∞ constraint (Eq. 94), the covariance equation (106) is modified as follows:

$$\tilde{A}\tilde{Q}_1 + \tilde{Q}_1\tilde{A}^T + r^{-2}\tilde{Q}_1\tilde{R}\ \tilde{Q}_1 + \tilde{V} = 0 \qquad \dots(109)$$

Relationship between original minimization problem and auxiliary minimization problem:

The original minimization problem

$$J(A_m,\ B_m,\ C_m) = \lim_{t\to\infty} E\left\{(y - y_m)^T\ R(y - y_m)\right\} \qquad \ldots(110)$$

s equivalent to minimization of

$$J(A_m,\ B_m,\ C_m) = \text{tr}\ \tilde{Q}\tilde{R} \qquad \ldots(111)$$

Subject to

$$\tilde{A}\tilde{Q} + \tilde{Q}\tilde{A}^T + \tilde{V} = 0 \qquad \ldots(112)$$

In order to accommodate H_∞ constraint, an auxiliary minimization problem was formulated and is given by

$$J\left(A_m,\ B_m,\ C_m\tilde{Q}_1\right) = \text{tr}\ \tilde{Q}\tilde{R} \qquad \ldots(113)$$

Subject to

$$\tilde{A}\tilde{Q}_1 + \tilde{Q}_1\tilde{A}^T + r^{-2}\ \tilde{Q}_1\tilde{R}\ \tilde{Q}_1 + \tilde{V} = 0 \qquad \ldots(114)$$

The results of L_2/H_∞ model reduction method are summarized in the following theorem:

<u>Theorem</u>

Suppose there exist nonnegative definite matrices $\hat{Q}$, $\hat{P}$, Q and P such that the following equations are satisfied.

$$AQ + QA^T + \gamma^{-2}Q\ \overline{\Sigma}\ Q + \tau_\perp\ \Sigma\ \tau_\perp^T = 0 \qquad \ldots(115)$$

$$A^T P + PA - \gamma^{-4}S^T PQ\ \Sigma\ QPS$$
$$+ \tau_\perp^T (I + \gamma^{-2}QPS)^T \overline{\Sigma}(I + \gamma^{-2}QPS)\tau_\perp = 0 \qquad \ldots(116)$$

$$\left(A - \gamma^{-4}Q\ \overline{\Sigma}\ QPS\right)\hat{Q} + \hat{Q}\left(A - \gamma^{-4}Q\overline{\Sigma}QPS\right)^T$$
$$+ \gamma^{-6}\hat{Q}S^T PQ\overline{\Sigma}QPS\hat{Q} + \Sigma - \tau_\perp\Sigma\tau_\perp^T = 0 \qquad \ldots(117)$$

$$\left(A + \gamma^{-2}Q\ \overline{\Sigma}\ \right)^T\hat{P} + \hat{P}\left(A + \gamma^{-2}Q\overline{\Sigma}\right) + (I + \gamma^{-2}QPS)^T\overline{\Sigma}\ (I + \gamma^{-2}QPS)$$
$$- \tau_\perp^T(I + \gamma^{-2}QPS)^T\overline{\Sigma}(I + \gamma^{-2}QPS)\tau_\perp = 0 \qquad \ldots(118)$$

$$\text{rank}\ \hat{Q} = \text{rank}\ \hat{P} = \text{rank}\ \hat{Q}\hat{P} = n_m \qquad \ldots(119)$$

$$\tau_\perp = i - \tau\ ,\ \tau = G^T\ \Gamma\ ,\ S = \left(I + \gamma^{-2}\hat{Q}P\right)^{-1} \qquad \ldots(120)$$

Then the reduced order model matrices are given by

$$A_m = \Gamma\left(A - \gamma^{-4} Q \overline{\Sigma} QPS\right) G^T \qquad \ldots(121)$$

$$B_m = \Gamma B \qquad \ldots(122)$$

$$C_m = C\left(I + \gamma^{-2} QPS\right) G^T \qquad \ldots(123)$$

Furthermore, the minimum value of auxilary case is given by

$$J_{min}, \; \hat{Q}_1 = tr \; \overline{\Sigma} \left(Q + \gamma^{-4} QPS \, \hat{Q} \, S^T P \, Q\right) \qquad \ldots(124)$$

V.   SUMMARY OF LQG/LTR DESIGN PROCEDURE

Among the various design methods for robust controllers the linear quadratic Gaussian with loop transfer recovery (LQG/LTR) design procedure has many advantages. This methodology is particularly suitable for the large flexibel structures. The LQG/LTR design procedure for system with minimum zeros with the loop broken at output can be summarized as follows:

(i)   Determine a nominal state variable model of the structure from NASTRAN finite element model. Augment pure integrators to the plant for zero steady state error.

(ii)   Define the desired performance loop shape by singular value plots.

(iii)   Design a Kalman filter with state noise covariance matrix $Q_e = \Gamma\Gamma^T$ and sensor noise matrix $Re = \mu I$. Choose the matrix Γ to meet the robustness characteristics at high and low frequencies.

(iv)   Design a linear quadratic regulator for recovering the stability margins of the closed loop system. The regulator weighting matrices are $Q_r = q^2 C^T C$ and $R_r = \rho I$. Plot singular value plots of $G(s)K(s)$ and vary q^2 until a desirable level of recovery is achieved.

(v)   Plot singular value plots of sensitivity function $[1 + G(s)K(s)]^{-1}$.

VI.   COMPARISON OF EIGENVALUES AND ZEROS OF REDUCED ORDER MODELS.

A large number of procedures are available in the literature for deriving reduced order models. The reduced order models for experimental grid are derived using balance-truncation method due to Moore[3], modified balance method by Prakash and Rao[15] and optimal projection [4] as method. For the sake of comparison purposes the eigenvalues and zeros of original system and reduced order models are given in Table 2 and 3 respectively. From the Table 3, it can be noticed that the number of nonminimum phase zeros are increased in reduced order models.

Table 2. Comparison of Eigenvalues

Original System		Balance-Truncation		Optimal Projection	
-1.2572D-02	+5.0289D+00j	-3.1330D+02	+1.2532D+01j	-3.1331D-02	+1.2532D+01j
-1.2572D-02	-5.0289D+00j	-3.1330D-02	-1.2532D+01j	-3.1331D-02	-1.2532D+01j
-3.1331D-02	+1.2532D+01j	-6.7691D-02	+2.7077D+01j	-6.7690D-02	+2.7077D+01j
-3.1331D-02	-1.2532D+01j	-6.7691D-02	-2.7077D+01j	-6.7690D-02	-2.7077D+01j
-6.7692D-02	+2.7077D+01j	-1.0441D-01	+4.1769D+01j	-1.0441D-01	+4.1769D+01j
-6.7692D-02	-2.7077D+01j	-1.0441D-01	-4.1769D+01j	-1.0441D-01	-4.1769D+01j
-8.0098D-02	+3.2039D+01j	-1.7859D-01	+7.1441D+01j	-1.7860D-01	+7.1441D+01j
-8.0098D-02	-3.2039D+01j	-1.7859D-01	-7.1441D+01j	-1.7860D-01	-7.1441D+01j
-1.0442D-01	+4.1768D+01j	-1.7952D-01	+7.2468D+01j	-1.7968D-01	+7.2469D+01j
-1.0442D-01	-4.1768D+01j	-1.7952D-01	-7.2468D+01j	-1.7968D-01	-7.2469D+01j
-1.7573D-01	+7.0291D+01j	-2.2569D-01	+9.0304D+01j	-2.2570D-01	+9.0303D+01j
-1.7573D-01	-7.0291D+01j	-2.2569D-01	-9.0304D+01j	-2.2570D-01	-9.0303D+01j
-1.7860D-01	+7.1441D+01j				
-1.7860D-01	-7.1441D+01j				
-1.8118D-01	+7.2474D+01j				
-1.8118D-01	-7.2474D+01j				
-2.2575D-01	+9.0302D+01j				
-2.2575D-01	-9.0302D+01j				
-2.6425D-01	+1.0570D+02j				
-2.6425D-01	-1.0570D+02j				

Table 3. Comparison of Transmission Zeros

Original System		Balance-Truncation Method				Optimal Projection Method	
		Moore		Modified			
-3.0658D-01	+1.1651D+02j	-1.1005D+04	+0.0000D+00j	-1.4167D+01	+6.1945D+01j	-1.1539D+04	+0.0000D+00j
-3.0658D-01	-1.1651D+02j	4.0082D+03	+0.0000D+00j	-1.4167D+01	-6.1945D+01j	4.7587D+03	+0.0000D+00j
1.1378D+01	+6.3812D+01j	-1.9519D+03	+0.0000D+00j	2.1391D-01	+6.3903D+01j	-2.1401D+03	+0.0000D+00j
1.1378D+01	-6.3812D+01j	1.5335D-01	+4.3353D+01j	2.1391D-01	-6.3903D+01j	1.2655D-01	+4.3344D+01j
-1.1665D+01	+6.3735D+01j	1.5335D-01	-4.3353D+01j	1.3560D+01	+6.1768D+01j	1.2655D-01	-4.3344D+01j
-1.1665D+01	-6.3735D+01j	1.6452D-01	+3.8622D+01j	1.3560D+01	-6.1768D+01j	1.3399D-01	+3.8637D+01j
-1.7505D-01	+6.9605D+01j	1.6452D-01	-3.8622D+01j	-1.3128D-02	+1.1245D+01j	1.3399D-01	-3.8637D+01j
-1.7505D-01	-6.9605D+01j	-1.1860D-01	+2.0987D+01j	-1.3128D-02	-1.1245D+01j	-1.2409D-01	+2.0981D+01j
-1.5554D-01	+6.6389D+01j	-1.1860D-01	-2.0987D+01j	6.4199D-01	+4.0325D+00j	-1.2409D-01	-2.0981D+01j
-1.5554D-01	-6.6389D+01j			6.4199D-01	-4.0325D+00j		
-8.0524D-02	+3.2195D+01j			-6.5164D-01	+3.9826D+00j		
-8.0524D-02	-3.2195D+01j			-6.5164D-01	-3.9826D+00j		
-2.5780D-02	+1.5495D+01						
-2.5780D-02	-1.5495D+01j						
6.8264D-01	+3.9936D+00j						
6.8264D-01	-3.9936D+00j						
-6.8329D-01	+3.9942D+00j						
-6.8329D-01	-3.9942D+00j						
-1.4988D-03	+3.3833D+00j						
-1.4988D-03	-3.3833D+00j						

47-24

VII. EFFECTS OF NON-MINIMUM PHASE ZEROS

One limitation of the LQG/LTR design procedure (Section V) is that it can obtain arbitrarily good recovery only for minimal phase plants. If the plant is non-minimum phase, the LTR technique can not recover the state feedback loop arbitrarily well. The original 10-mode state variable model of the experimental grid structure has 3-pairs of non-minimum phase zeros. Hence the effects of NMP zeros on the closed loop system performance and techniques for loop transfer recovery for system with NMP zeros is investigated in this report.

The objective of a LTR procedure is to recover guaranteed stability margins by selecting proper linear quadratic regulaor (LQR) gains

$$Q_r = H^T H + q^2 C^T C \qquad\qquad ...(125)$$
$$R_r = pI \qquad\qquad ...(126)$$

As $q^2 \to \infty$, the eigenvalues of closed loop system will move towards the plant transmission zeros. Since $q^2 C^T C$ is at least positive semidefinite, the closed loop eigenvalues will never lie in the right half plane for any value of q^2. Hence the closed loop engenvalues do not move towards the non-minimal phase zeros as $q^2 \to \infty$. Unlike the minimal phase case, we can not see any pattern on the poles and zeros of the compensator for various values of q^2. The effects of non-minimal phase zeros on the closed loop system performance can be summarized as follows:

(a)   If the NMP zeros lie well outside the required bandwidth, the recovery procedure will give satisfactory performance. If any of them lie within the required bandwidth, the closed loop system may become unstable near NMP zero frequencies

(b)   If the plant is minimum phase the recovery procedure will cancel all undesired dynamics of the plant i.e., as $q \to \infty$, $K(s) \to G^{-1}(s)C\Phi(s)K_f$. In the case of non-minimum phase systems, the recovery process can not cancel all undesired dynamics. These unwanted dynamics will tend to aggravate the compensator.

This phenomemon is observed in the grid structure at 13.85 Hz (87 rad/sec). The singular value plot of the target feedback loop and G(s)K(s) is shown in Figure 3. The sensitivity plots are given in Figure 4.

(c)   If the recovery procedure introduces additional poles in the vicinity of the desired closed loop poles, an additional gain and a phase shift of (-180°) is introduced in the loop and closed loop system will be unstable at that frequency. This phenomenon has been observed at a frequency 13.85 Hz for a value of $q^2 = 500$.

VIII. EXPERIMENTAL RESULTS

A block diagram representation of experimental setup is given in Figure 5. The accelerometer were mounted at mode number 14, 21, and 27 on the grid. The x-axis torquers are located at mode numbers 22 and 24 and y-axis torquer is placed at node 18 on the grid. The Max 100 computer is used for the implementation of the controller. The Max 100 computer was loaded with real time code generated using Matrixx (software package) capabilities.
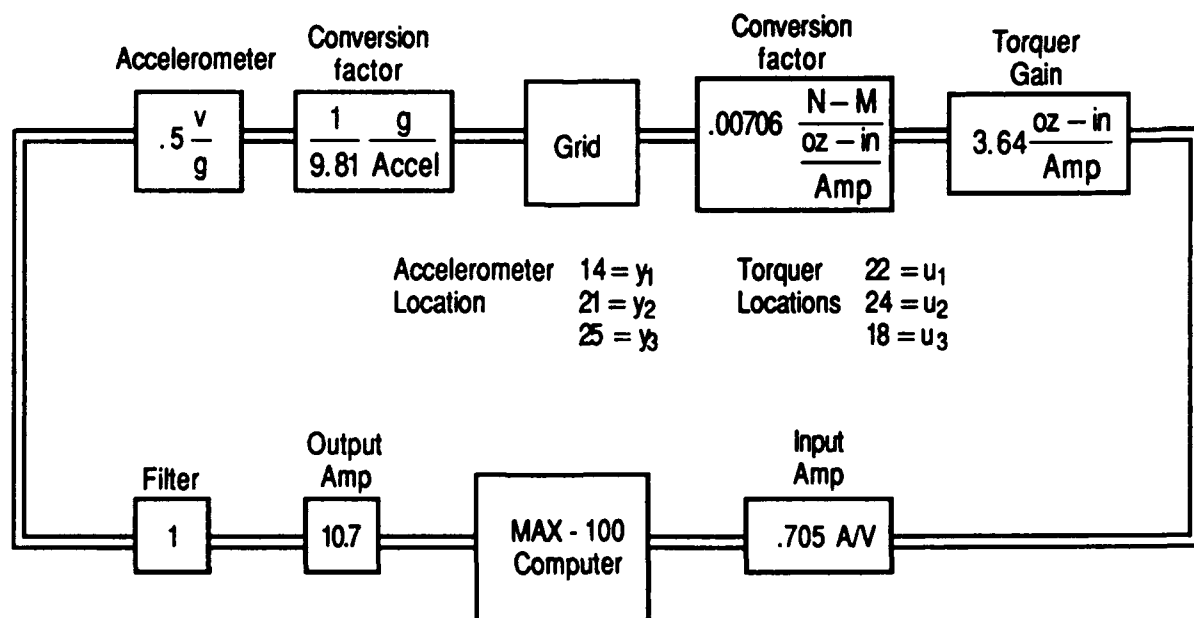
Figure 5. Block Diagram Representation of Experimental Setup

Several LQG/LTR controllers were designed and implemented on an overlapped experimental grid structure. A 20th order model of the grid is

determined by using NASTRAN finite element model. To minimize the computational and implementation requirements of the controller, a 12th order reduced model is derived by using balance-truncate procedure.

To minimize the computational requirements, a linear transformation is performed on the LQG/LTR compensator to convert into modal form. A discrete equivalent of this controller is obtained with a sampling frequency of 80 Hz.

The 12th order LQG/LTR controller exhibited high loop gain at mode #9 (13-85 Hz). In addition a (-180°) phase shift in the controller output is observed and the closed loop system was unstable at that mode.

We repeated the controller design procedure and implementation using 10th order, 8th order and 6th order models. In the 10th order compensator system, the mode at 11.5 Hz become unstable. In the 8th order system the mode at 11.05 has become marginally stable. The 6th order compensator gave stable performance at all modes.

Open and closed loop responses with 6th order controller for mode #5 is given in Figure 6. To illustrate the robustness properties of the controller, the following plant perturbations were performed:

Perturbation #1: 70g masses were removed from structure nodes 7 and 9.

Perturbation #2: Two more 70g masses were removed from structure nodes 16 and 20.

Perturbation #3: A mass of 278g is added at node 17 of the structure.

Perturbation #4: A mass of 278g is added at node 11 (flexing mode) of the structure.

The structural mode shape for 5th mode is given in Figure 7. The initial condition reponses with these perturbations are presented in Figures 8 and 9.

Freq in rad per sec
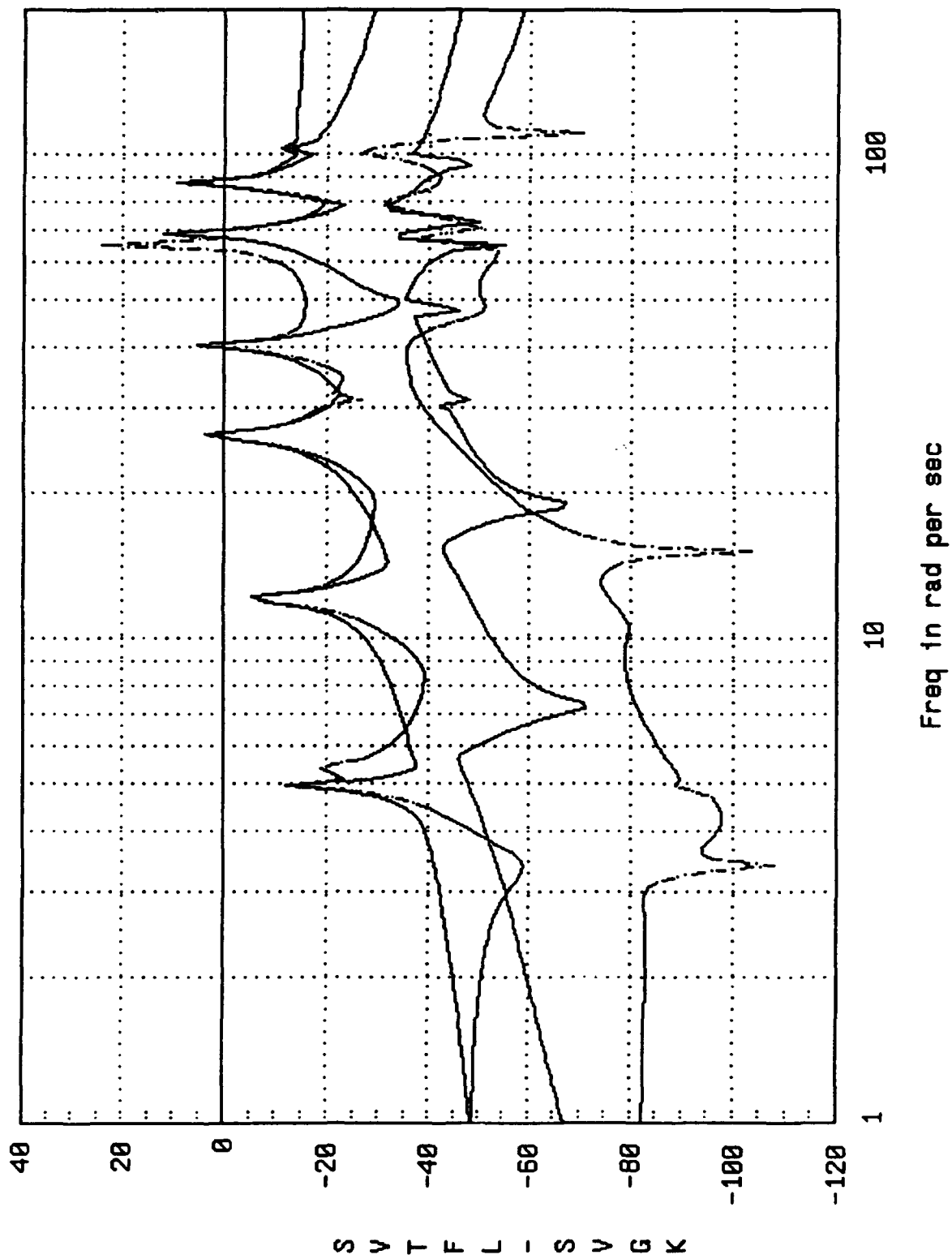
Fig. 3. Singular Value Plots of TFL and G(s)K(s).
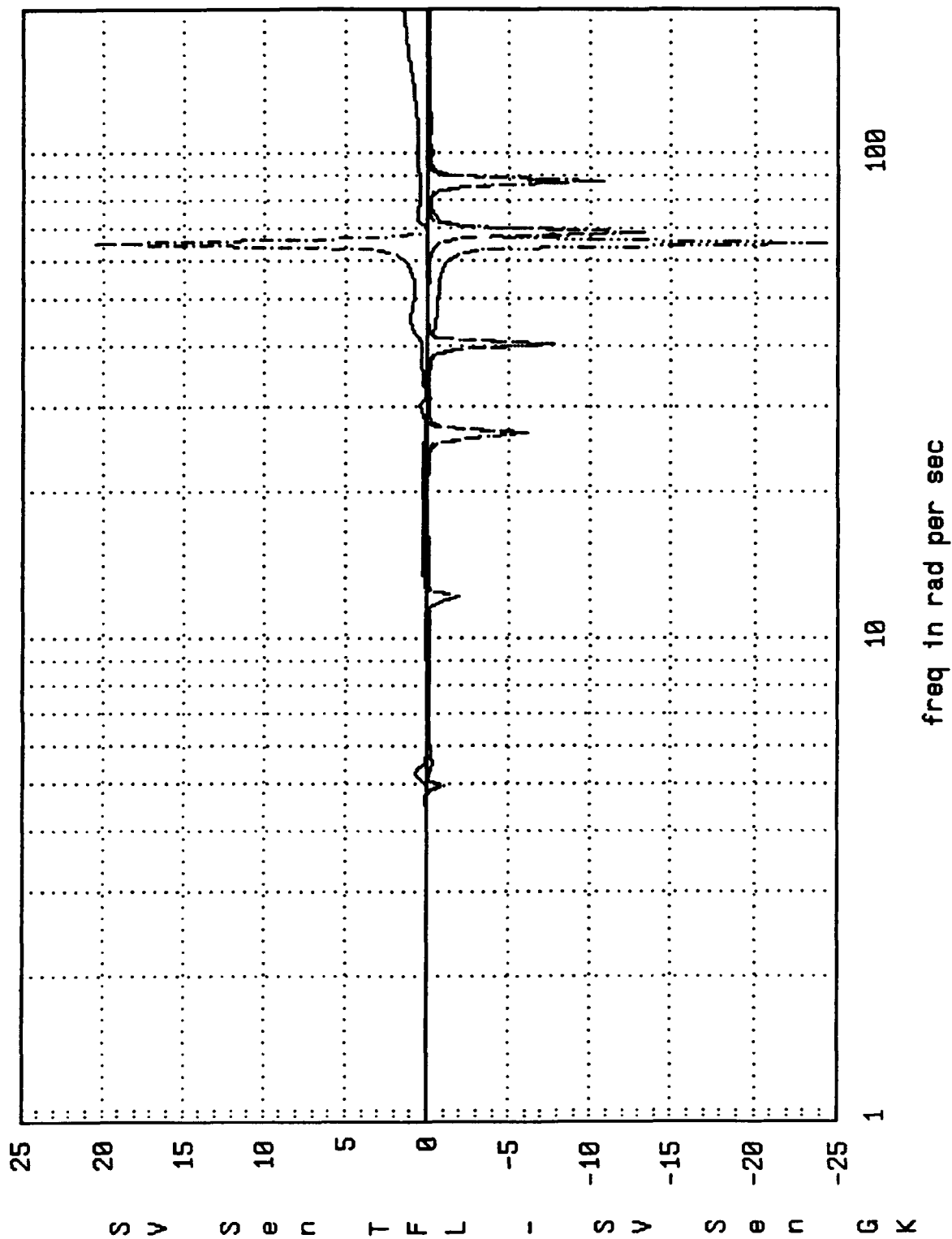
Fig. 4. Sensitivity Plots of TFL and G(s)K(s).
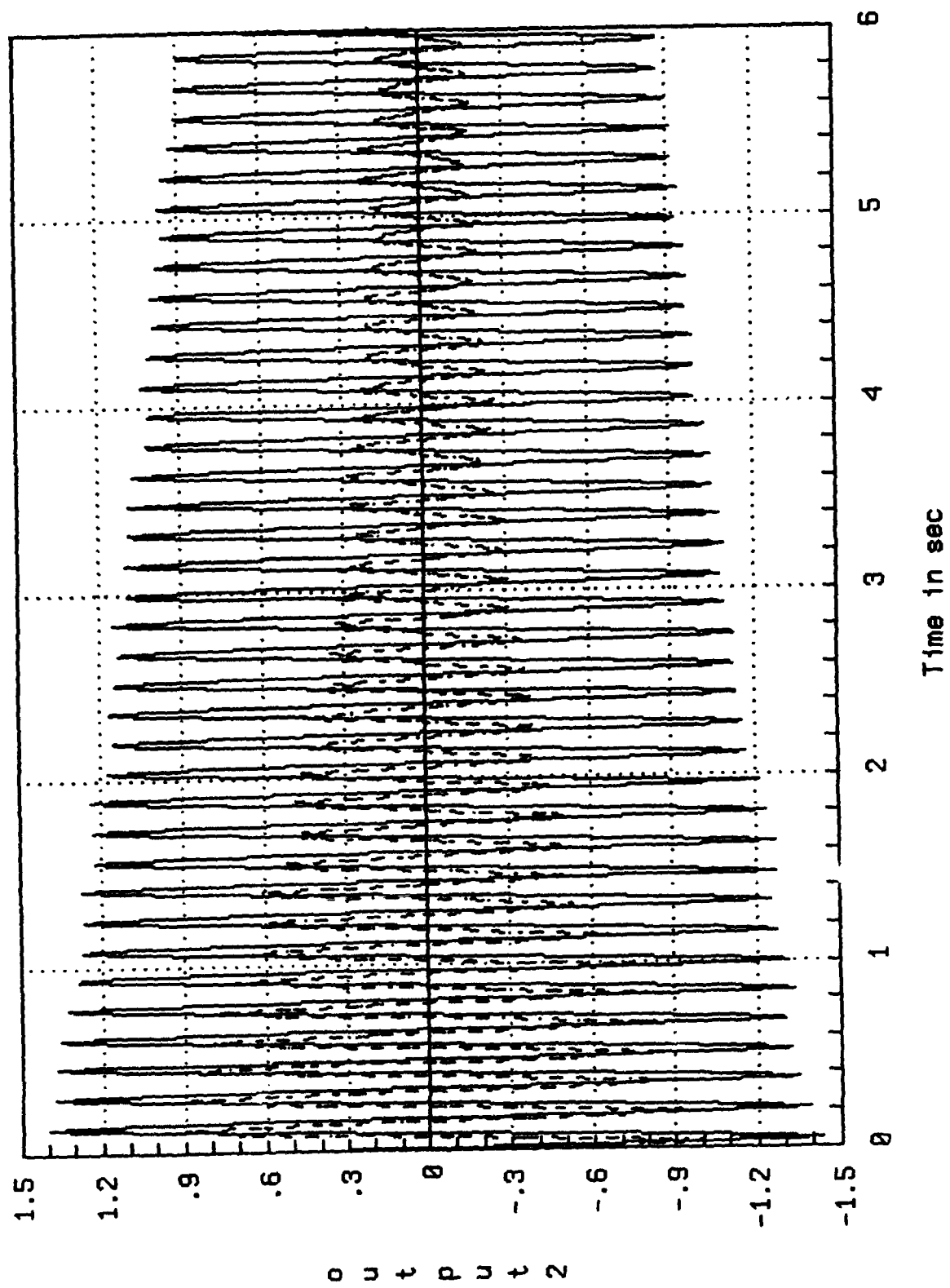
Fig. 6. Open and Closed Loop Response of Mode #5.

7/ 5/84   MAX-DEF. • 1.09142620
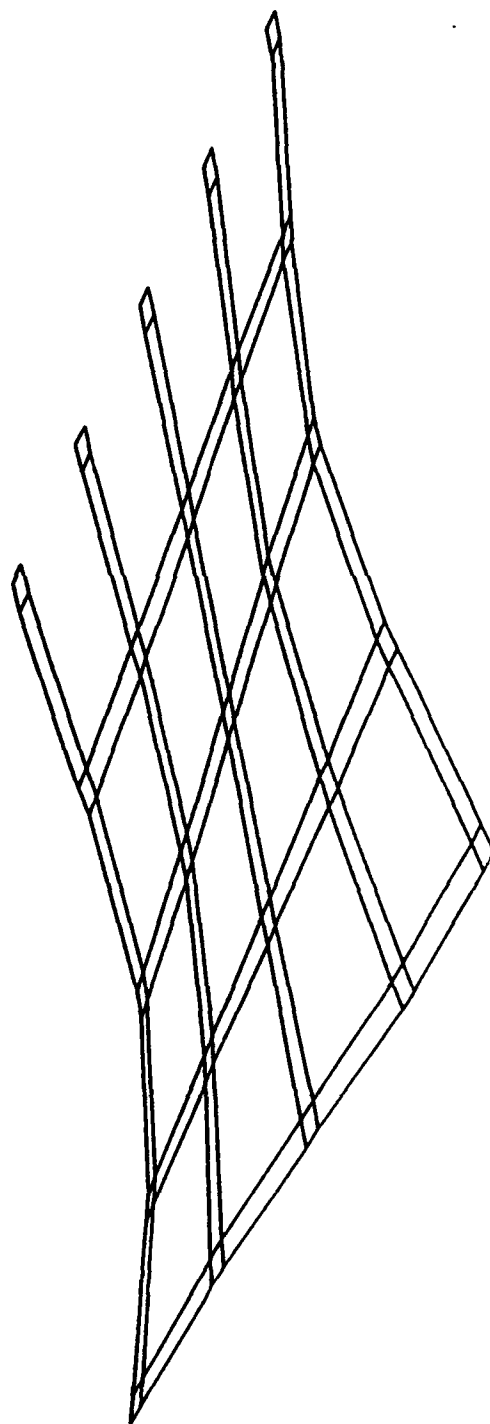
Fig. 7. The Structural Model Shape for 5th Mode.
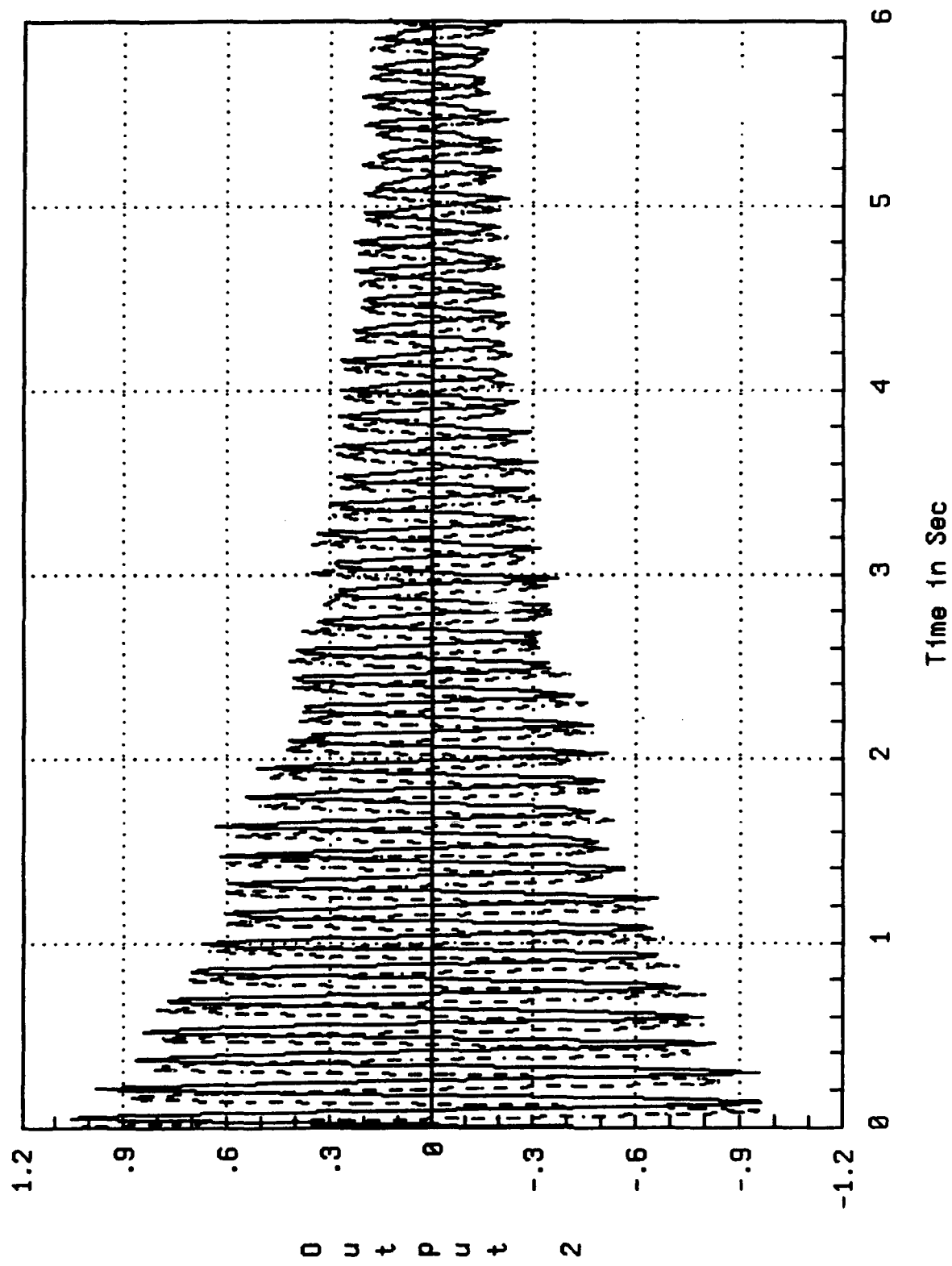
Fig. 8. Closed Loop Response of Mode #5 with Perturbations 1 and 2.
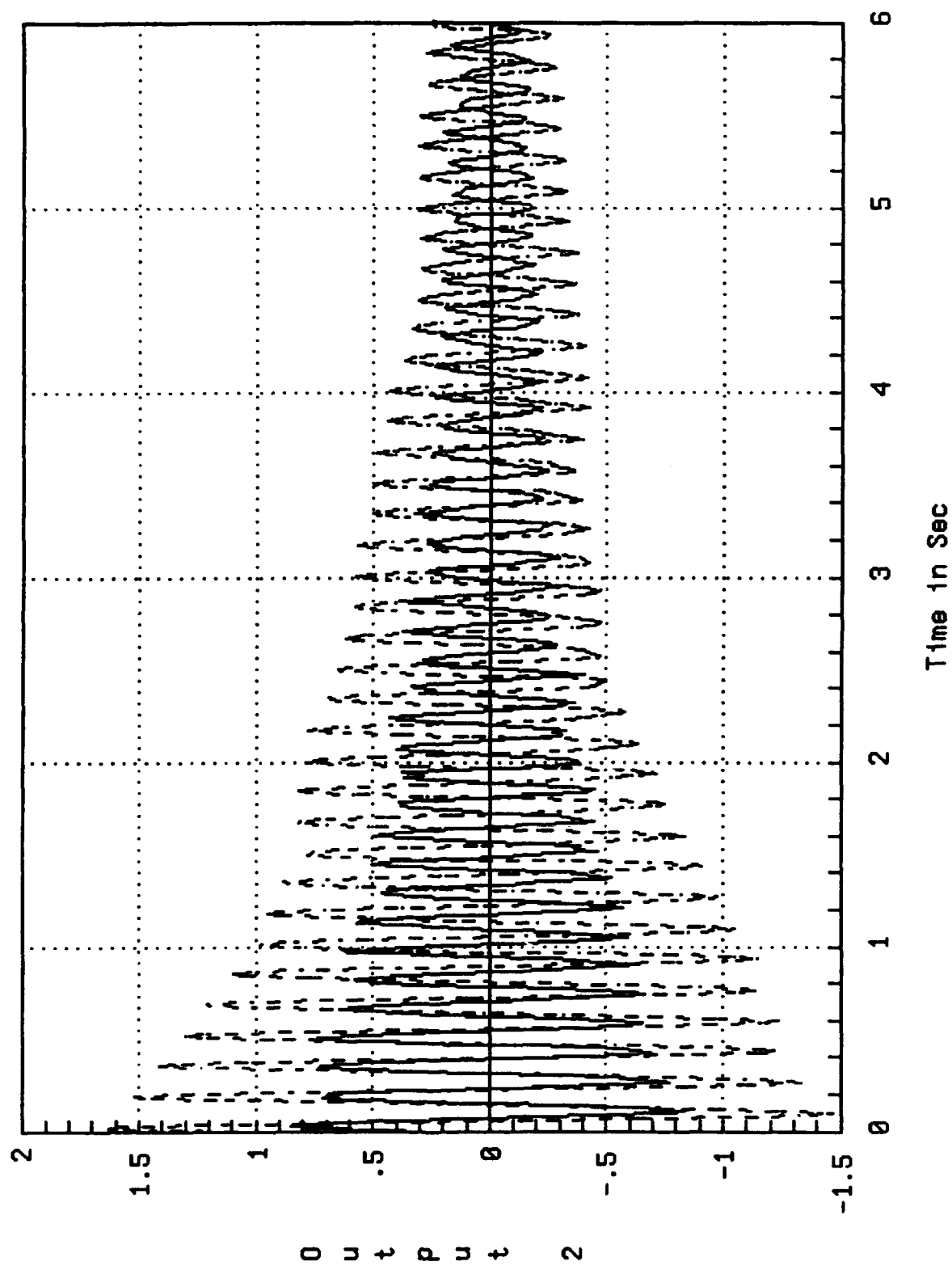
Fig. 9. Closed Loop Response of Mode #5 with Perturbations 3 and 4.

IX. CONCLUSIONS AND RECOMMENDATIONS

Robust control design methodologies were tested on an experimental grid structure. The balance-truncation, optimal projection and combined L_2/H_∞ model reduction methods were employed to derive reduced order models. The numerical algorithms are developed for optimal projection methods.

The reduced order LQG/LTR controllers were implemented on the experimental grid structure using real time computers. The reduced order models of the grid have non-minimum phase zeros and the presence of these zeros aggravated the controller action at certain frequencies. The effects of non-minimum phase zeros on the closed loop system were identified and reported in this report. The closed loop system performance under various plant perturbation was experimentally determined and found satisfactory.

In the present study, the dynamics of actuators and sensors are not considered. It is recommended that these dynamics be augmented with the dynamics of grid structure. It is also recommended to design H_∞ optimal controllers for grid structure and compare with robust controllers.

X. ACKNOWLEDGEMENTS

XI REFERENCES

[1]   S.S. Lamba and S. Vittal Rao, "On Suboptimal Control Via the Simplified Models of Davison, "*IEEE Trans. on Automatic Control*, Vol. AC-19, No-4, 1974, pp 448-450.

[2]   R.E. Skelton, P. Hughes and H. Hablani, "Order Reduction for Models of Space Structures Using Modal Cost Analysis," *AIAA J. of Guidance, Control and Dynamics*, Vol. 5, No. 4, July 1982.

[3]   B.C. Moore, "Principal Component Analysis in Linear Systems: Controllability, Observability and Model Reduction," *IEEE Trans. on Automatic Control*, Vol, AC-26, 1981, pp 17.32.

[4]   D.C. Hyland and D.S. Bernstein, "The Optimal Projection Equations for Fixed-Order Dynamic Compensation," *IEEE Trans. on Automatic Control*, Vol. AC-29, No.-11, Nov. 1984, pp 1034-1037.

[5]   ᴸ.C. Hyland and D.S. Bernstein, "The Optimal Projection Equations for Model Reduction and the Relationships Among the Methods of Wilson, Skelton and Moore," *IEEE Trans. on Automatic Control*, Vol. AC-30, No. 12, Dec. 1985, pp 1201-1211.

[6]   W.M. Haddad and D.S. Bernstein, "Combined L_2/H_∞ Model Reduction," *Int. J. of Control*, Vol. 49, No. 5, pp 1523-1435, 1989.

[7]   N. Sundararajan, S.M. Joshi and E.M. Armstrong, "Robust Controller Synthesis for a Large Flexible Space Antenna," *AIAA J. of Guidance, Control and Dynamics*, Vol. 10, No. 2, 1987, pp 201-208.

[8]   S.M. Joshi and E.S. Armstrong, "Design of Robust Line-of-Sight Pointing for t ᵊ SCOLE Configuration," Proceedings of American Control Confᴜᵣence, 1987, pp 1125-1127.

[9]   R.K. Yadavalli, "Robustness Stabilization Under Mode Truncation and
      Parameter Variations," Proceedings of American Control Conference,
      1987, pp 490-495.

[10] D.C. Hyland and D.S. Bernstein, "The Optimal Projection Equations for
      Model Reduction and the Relationships Among the Methods of Wilson,
      Skelton and Moore," *IEEE Trans. on Automatic Control*, Vol. AC-30, pp
      1201-1211, 1988.

[11] A. Das, T.J. Strange, W.T. Schlaegel and J. M. Ward, "Experiments in
      Modelling and Parameter Estimation of Flexible Structures,"
      Proceedings of 2nd NASA/DOD, CSI Technology Conference, Nov.
      1987.

[12] M.A. Norris, R.C. Thomson and A. Das, "Low-Frequency Response of
      Accelerometer for Observer Design in a Gravity Environment," *AIAA J.
      of Guidance, Control and Dynamics*, 1990.

[13] K.T. Ngo, "Design of Reduced Order Robust Controller Using Optimal
      Projects on Method," M.S. Thesis, University of Missouri-Rolla.

[14] S. Richter and E.G. Collins, Jr., "A Homotopy Algorithm for Reduced
      Order Controller Design Using the Optimal Projection Equations,"
      Proceedings of the IEEE Conference on Decision and Control, Tampa,
      pp 506-511, Dec. 1989.

[15] R. Prakash and S. Vittal Rao, "Model Reduction by Low-Frequency
      Approximation of Internally Balanced Representation," Presented at the
      28th Control and Decision Conference, Dec. 1989.

1990 USAF-UES RESEARCH INITIATION PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.

FINAL REPORT

Theoretical Study of Capillary Pumping in Heat Pipes

Principle Investigator: Larry W. Swanson

Prepared by: Larry W. Swanson and
                             Gregory C. Herdt

Department: Department of Engineering

University: University of Denver

Contract Number: F49620-88-C-0053/SB5881-0378

ABSTRACT

This report addresses the development of two mathematical models which can be used to improve the pumping capability of capillary wick structures. The first model describes the evaporating meniscus in a capillary tube from both a macroscopic and microscopic point of view. The second model specifically focuses on a description of the microscopic physics near the interline, namely the van der Waals forces, and includes both curvature and retardation effects.

The model describing the evaporating meniscus in a capillary tube includes the full three-dimensional Young-Laplace equation, Marangoni convection, van der Waals dispersion forces, and nonequilibrium interface conditions. The results show that varying the dimensionless superheat has no apparent effect on the meniscus profile. However, varying the dispersion number produces a noticeable change in the meniscus profile, but only at the microscopic level near the tube wall. No change in the apparent contact angle is observed with changes in the dimensionless superheat or dispersion number. In all cases, the dimensionless mean curvature is asymptotic to a value equal to that for a hemispherical meniscus. The local interfacial mass flux and total mass transfer rate increase dramatically as the dispersion number is increased suggesting that surface coatings can play a important role in improving or degrading capillary pumping. The model also predicts that the local capillary pressure remains constant and equal to $2\sigma/r_c$ regardless of changes in the dimensionless superheat and dispersion number.

ACKNOWLEDGEMENT

TABLE OF CONTENTS

1. Introduction

2. Objectives

3. Model of the Evaporating Meniscus in a Capillary Tube

    3.1. Literature Survey
    3.2. Formulation of Mathematical Model
    3.3. Solution Method
    3.4. Results/Discussion
    3.5. Cor.clusions
    3.6. Nomenclature

4. Van der Waals Forces in Capillary Tubes

    4.1. Introduction

        4.1.1. Motivation for the Study
        4.1.2. Historical Background

    4.2. Theory of Van der Waals Forces in a Capillary Tube

        4.2.1. Assumptions of the Theory
        4.2.2. Surface Modes in a Capillary Tube
        4.2.3. Calculation of the Free Energ

    4.3. Discussion

        4.3.1. Models of Dielectric Behavior
        4.3.2. The Computer Program

    4.4. Results and Conclusions

        4.4.1. Summary
        4.4.2. Remarks on the Results
        4.4.3. Conclusions

    4.5. Nomenclature

5. Recommendations

6. References

7. Appendix A (Functions for chapter 3)

8. Appendix B (Computer code for chapter 3)

9. Appendix C (Lifshitz theory)

10. Appendix D (Computer code for chapter 4)

11. Appendix E (Optical data files for chapter 4)

12. Appendix F (Results for chapter 4)

1. INTRODUCTION

The capillary forces present in the evaporator and the condenser sections of capillary heat pumps (e.g. heat pipes or capillary pumped loops) constitute the basic driving force for internal fluid circulation. Under normal operating conditions, the rate at which fluid is circulated determines energy transport capacity of the heat pump. The net capillary force is generated by the integral effect of the evaporating and condensing menisci.

Despite the complicated and interactive nature of these capillary forces, the heat transport capability of capillary heat pumps is usually determined by equating the capillary pressure to the liquid pressure drop in the wick or loop, assuming that the adverse capillary pressure in the condenser is negligible and the capillary pressure in the evaporator is maximum. The capillary pressure is commonly taken as $2\sigma/r_c$, where σ is the vapor-liquid surface tension and r_c is the average capillary radius. This is known as the wicking limitation and presumably represents the point at which dry-out begins in the evaporator. In most situations, the maximum capillary pressure can not be used to evaluate the local interfacial mass transfer rate which is an intermediate boundary condition between the vapor and liquid phases. This is due to the fact that the transport phenomena in the thin film on the capillary wall, which is coupled to the meniscus, is dramatically affected by either evaporation or condensation. Under certain conditions, the interfacial processes may even be rate limiting.

Although capillary heat pumps have been studied extensively in the past, very little effort has been put forth to understand the basic mechanism describing the evaporating and condensing menisci in capillary structures. For example, past studies which address the vapor-liquid

interface in heat pipes (Ambrose el al. [1987], Beam [1985], Beam [1985]) treat the interface as a macroscopic phenomenon and are semiempirical in nature since they require mass transfer data. This approach is very interesting and useful for predicting dry-out in certain capillary structures; however, a mathematical model derived from first principles should actually generate values for the mass transfer rate since the mass transfer rate is not a true parameter of the system. A second drawback of past macroscopic models describing the vapor-liquid interface is that they neglect microscopic interfacial phenomena taking place near the vapor-liquid-solid contact line, known as the interline. The pioneering work of Wayner and his coworkers [1976, 1976, 1979, 1986] has shown that these effects significantly influence both the apparent contact angle and the heat flux in an evaporating meniscus for a smooth plane wall as well as a capillary tube. The inclusion of these effects into an integral formulation of the macroscopic interface similar to that derived in Ambrose et al. [1987] will yield a solution that does not require the use of mass transfer data. In capillary heat pumps, this type of integral formulation will also be amenable to formal optimization techniques which can be used to determine the optimal radial and axial pore size distributions in the capillary structure. These distributions represent fabrication parameters which are essential to the design of advanced capillary structures. The foundation of this design technique, however, must begin with a fundamental understanding of the evaporating and condensing meniscus, preferably for a simple geometry characterizing common wicking material. A capillary tube configuration was chosen for this study because it represents a pore geometry both compliant to formal mathematical characterization and common to capillary structures.

2. OBJECTIVES

The primary objective of this work was to improve our understanding of the evaporating meniscus in a capillary tube in order to identify what types of modifications should be made to capillary structures to enhance their heat transport capability. To satisfy this objective, we initially formulated and solved a mathematical model of the evaporating meniscus in a capillary tube to determine how the various forces at the interface influence the meniscus profile, the mass transfer rate, and the capillary pressure. The model utilized a very crude approximation to the van der Waals forces near the triple interline and indicated that a more thorough understanding of the van der Waals forces in this region was necessary. Therefore, we developed an extensive theoretical model of the van der Waals forces near the triple interline incorporating both retardation and curvature effects. The results from the microscopic model will help determine the role that wick curvature and fluid retardation play in the microscopic forces, which in turn affects the pumping capability of capillary structures.

3. MODEL OF THE EVAPORATING MENISCUS IN A CAPILLARY TUBE

3.1. Literature Survey

Studies specifically addressing evaporation and condensation in capillary tubes have been undertaken by a number of investigators. To our knowledge, the earliest study of evaporation in isolated capillary tubes was conducted by Derjaguin et al. [1965]. They developed a model which included an adsorption isotherm based on dispersion forces as well as mass transport due to both vapor diffusion and film flow. They found that including mass transport in the thin film near the interline dramatically increased the rate of evaporation over that estimated solely by diffusion theory. They also confirmed their theoretical results with experimental data.

Preiss and Wayner [1976] conducted an experimental study of the evaporating meniscus near the exit of a capillary tube with ethanol as the working fluid. The meniscus was found to be stable over a wide range of evaporation rates and hydrostatic heads. They also concluded that fluid flow to the base of the stable evaporating meniscus was caused by a change in its curvature. Experiments at high evaporation rates demonstrated that sputtering occurred near the wall before the meniscus became unstable.

Wayner [1979] also developed a dispersion-based model of the evaporating meniscus near the exit of a capillary tube. He found that viscous flow in the thin film near the interline significantly affected the entire meniscus profile. The model also predicted that increasing the interline heat flux reduced the capillary pressure by increasing the apparent contact angle. The impact of the dispersion forces on the evaporation process was highlighted by a comparison of the interline heat sink capability for

diffe ent liquid-solid substrate combinations.

Philip [1977] performed a theoretical study of condensation in capillary pores. His model included dispersion forces, a cylindrical representation of the surface tension forces, and assumed that the capillary pressure was constant over the entire meniscus. An iterative method was used to determine the capillary pressure and the condensate film thickness downstream from the meniscus. He found that his solution produced capillary pressures roughly 44% higher than those predicted by the classical Kelvin solution.

Finally, Meyer [1984] performed very interesting analytical assessments of evaporation in 1) isolated capillaries, 2) two interconnected capillaries of variable cross-section, and 3) networks of capillary menisci. He found that large temperature gradients developed in localized neighborhoods of the menisci. For interconnected menisci, evaporation was found to proceed in statically unstable configurations, via so-called Haines jumps (Haines [1930]), due to the dynamic balance of surface tension, local evaporation rate, and viscous shear. His model unfortunately did not include dispersion forces common to the previous analytical studies.

The remainder of Chapter 3 is dedicated to the development of a more rigorous model of the evaporating meniscus in a capillary tube. The results will be used to identify what types of modifications should be made to capillary structures to enhance their heat transport capability.

3.2. Formulation of the Mathematical Model

The forthcoming mathematical model describes the fluid mechanics, heat transfer, and interfacial phenomena characteristic of an isolated evaporating meniscus in a

capillary tube. The formulation includes the full three-dimensional Young-Laplace equation, Marangoni convection, van der Waals dispersion forces, and nonequilibrium interface conditions. The formulation of this model has been strongly influenced by the work of Wayner and his coworkers (cited earlier) as well as the study of Mirzamoghadam and Catton [1988].

Figure 3.1 shows the various flow regions characterizing the transport processes near the evaporating meniscus in a capillary tube. The origin of the axial coordinate (x) is at the interline while the origin of the radial coordinate (r) is at the tube center. r_i is the distance from the axial centerline to the meniscus interface. Each region is briefly described as follows:

  1) Thin film region (< 1 micron) - the van der Waals forces (intermolecular dispersion forces) are important and significantly affect the transport phenomena in the thin film.

  2) Meniscus region (> 1 micron) - the van der Waals forces are small and the interfacial curvature dominates the capillary pressure creating an attenuated Hagen-Poiseuille flow field.

  3) Hagen-Poiseuille region - Hagen-Poiseuille flow.

A mathematical model of the process can be formulated under
the following assumptions:

1) Axial symmetry.
2) Steady-state two-dimensional laminar flow.
3) Incompressible flow.
4) The advective and convective terms, with the exception
of axial convection, are negligible.
5) The radial pressure gradient is negligible.
6) The temperature in the bulk vapor is equal to the
interfacial vapor temperature.
7) The pressure in the bulk vapor is constant.
8) No slip at the wall.
9) Asperities on the wall do not affect any of the
underlying forces.
10) σ is not affected by interfacial curvature.
11) Retardation effects in $\overline{A}$ are negligible.
12) The hydrostatic pressure is negligible.
13) Marangoni effects are important only in the thin film
region.

Under these assumptions, the momentum equation describing
flow in the meniscus and interline regions is

$$-\frac{dP_u}{dx} + \frac{\mu_l}{r}\frac{\partial}{\partial r}\left(r\frac{\partial u_l}{\partial r}\right) = 0 ,$$

(3.1)

with boundary conditions of

$$r = r_i, \qquad -\mu_l\frac{\partial u_l}{\partial r} = \frac{\partial \sigma}{\partial x} ,$$

$$r = r_c, \qquad u_l = 0 .$$

The first boundary condition equates the interfacial shear
stress to the change in surface tension with respect to
position, thus accounting for Marangoni effects at the
interface. The solution to this differential equation is
straightforward yielding

$$u_l = \frac{1}{4\mu_l}\frac{dP_{li}}{dx}(r^2 - r_c^2) + \left(\frac{r_i}{\mu_l}\frac{d\sigma}{dx} - \frac{r_i^2}{2\mu_l}\frac{dP_{li}}{dx}\right)\ln\frac{r}{r_c}.$$
(3.2)

Applying l'Hospitals rule shows that this profile correctly reduces to the Hagen-Poiseuille velocity profile when the interface is at the axial center of the tube ($r = r_i = 0$). The liquid pressure in the velocity profile can be evaluated using a modified form of the Young-Laplace equation given by

$$P_{li} = P_{vi} - 2\sigma K + \frac{\overline{A}}{(r_c - r_i)^3}.$$
(3.3)

The second term on the right side of the equation is the capillary pressure due to interfacial curvature and the third term is a disjoining pressure (Deryagin et al. [1965]) for nonpolar liquids due to the van der Waals forces near the interline. $\overline{A}$ is the dispersion coefficient and is negative for a wetting fluid. Note that in the thin film near the interline ($r_i \approx r_c$), the disjoining pressure significantly influences the capillary pressure; the disjoining pressure also decreases rapidly as the interface gets further away from the tube wall. Equation (3.3) also assumes that the pressure forces, surface tension forces, and dispersion forces are simply additive at the vapor-liquid interface. This assumption is in accord with the thermodynamic (or force balance) derivation of the standard Young-Laplace equation for an isolated vapor-liquid interface found in Adamson [1982] and is probably accurate to leading order. However, it is important to note that the dispersion coefficient ($\overline{A}$) in the thin film is not only dependent on the dielectric behavior characterizing the liquid and vapor, but also the dielectric behavior of the solid substrate (e.g. see Israelachvili [1985]). The

surface tension for an isolated vapor-liquid interface can also be written in terms of the dispersion forces between the vapor and liquid (Ivanov [1988]); thus, a rigorous derivation of th capillary pressure $(P_{vi} - P_{li})$ from first principles should inherently account for solid-liquid-vapor interactions. A unified expression of this type, in effect, will attenuate the vapor-liquid surface tension in the thin film due to the presence of the solid substrate and also produce a disjoining pressure for thin films when there is no interfacial curvature.

An expression for the mean interfacial curvature in a capillary tube has been derived by Philip [1977] and is given as

$$K = \frac{1}{2}\left\{\frac{1}{r_i[1 + (dr_i/dx)^2]^{1/2}} - \frac{d^2 r_i/dx^2}{[1 + (dr_i/dx)^2]^{3/2}}\right\}. \qquad (3.4)$$

The mass flowing between the interface and wall (mass transfer rate) at any axial position in the liquid can be found by multiplying the liquid velocity by the liquid density and integrating over the flow cross-sectional area between the meniscus interface and the tube wall. This yields

$$\dot{m} = \frac{1}{F_1}\frac{dP_{li}}{dx} - \frac{F_2}{F_1}\frac{d\sigma}{dT_i}\frac{dT_i}{dx}, \qquad (3.5)$$

where the chain rule has been used in the second term to accentuate the Marangoni effect $(d\sigma/dT_i)$. The functions F_1 and F_2 are cumbersome functions of ρ_l, μ_l, r_c, and r_i, and are given in the Appendix A. The integral form of the mass continuity equation in the liquid is

$$r_i v_{li} = \frac{d}{dx} \int_{r_i}^{r_c} u_l \, r \, dr + r_i u_{li} \frac{dr_i}{dx},$$ (3.6)

where the Leibnitz rule has been used to generate the second term on the right hand side. Multiplying this expression by $2\pi\rho_l$, substituting for u_l, integrating and solving for v_{li} gives

$$v_{li} = \frac{1}{2\pi\rho_l r_i} \frac{d\dot{m}}{dx} + u_{li} \frac{dr_i}{dx}.$$ (3.7)

The interfacial velocity of the liquid normal to the interface is

$$w_{li} = v_{li} \cos\theta + u_{li} \sin\theta,$$ (3.8)

where θ is the local contact angle given by

$$\theta = \tan^{-1}\left(-\frac{dr_i}{dx}\right).$$ (3.9)

Equations (3.7), (3.8), and (3.9) can be combined to produce

$$w_{li} = \frac{\cos\theta}{2\pi\rho_l r_i} \frac{d\dot{m}}{dx},$$ (3.10)

which shows the relationship between the liquid velocity normal to the interface, the local contact angle, and the change in the mass transfer rate with respect to position. An expression for the nonequilibrium vapor mass flux at the interface has been derived by Schrage [1953] using kinetic

theory. Assuming that the change in temperature between the interface and the bulk vapor is small, this expression reduces to

$$-\rho_v w_{vi} = \left(\frac{2C}{2-C}\right)\left(\frac{1}{2\pi R T_i}\right)^{1/2} (P_{vi} - P_v),$$ (3.11)

where C is the accommodation coefficient. A mass balance at the interface requires that

$$\rho_v w_{vi} = \rho_l w_{li}.$$ (3.12)

Combining equations (3.10), (3.11), and (3.12) gives an expression for the interfacial pressure in the vapor phase,

$$P_{vi} = P_v - \left(\frac{2-C}{2C}\right)\left(\frac{RT_i}{2\pi}\right)^{1/2} \frac{\cos\theta}{r_i} \frac{d\dot{m}}{dx}.$$ (3.13)

Substituting this expression into equation (3.3) yields

$$P_{li} = P_v - \left(\frac{2C}{2-C}\right)\left(\frac{RT_i}{2\pi}\right)^{1/2} \frac{\cos\theta}{r_i} \frac{d\dot{m}}{dx} - 2\sigma K + \frac{\overline{A}}{(r_c - r_i)^3}.$$ (3.14)

Conservation of energy at the interface requires setting the heat flux normal to the interface equal to the latent heat required to evaporate the liquid, or after some manipulation

$$\frac{dT_i}{dx}\frac{dr_i}{dx} - \left.\left|\frac{\partial T}{\partial r}\right|\right._{r=r_i} = \frac{h_{fg}}{2\pi\rho_l r_i}\frac{d\dot{m}}{dx}.$$ (3.15)

Assuming axial diffusion is small, conservation of energy in the liquid can be expressed as

$$u \frac{\partial T}{\partial x} = \frac{\alpha}{r} \frac{\partial}{\partial r} \left(r \frac{\partial T}{\partial r} \right),$$

(3.16)

with boundary conditions of

$r = r_i,$ $T = T_i(x),$

$r = r_c,$ $T = T_w =$ constant.

The solution to this partial differential equation is approximated using the von Karman integral method assuming a linear temperature profile of

$$T = T_i + \frac{T_w - T_i}{r_c - r_i} (r - r_i).$$

(3.17)

The resulting expression is

$$\left(\frac{F_3}{4\mu_l} \frac{dT_i}{dx} - \frac{r_i^2 F_5}{2\mu_l} \frac{dT_i}{dx} + \frac{F_4}{4\mu_l} \frac{dr_i}{dx} - \frac{r_i^2 F_6}{2\mu_l} \frac{dr_i}{dx} \right) \frac{dP_{li}}{dx}$$
$$+ \left(\frac{r_i F_5}{\mu_l} \frac{d\sigma}{dT_i} \frac{dT_i}{dx} + \frac{r_i F_6}{\mu_l} \frac{d\sigma}{dT} \frac{dr_i}{dx} \right) \frac{dT_i}{dx} = \alpha(T_w - T_i),$$

(3.18)

where F_3, F_4, F_5, and F_6, are cumbersome functions of r_c, r_i, T_c, and T_i, and are given in the Appendix A. All equations to this point are now nondimensionalized using the following dimensionless variables:

$$\bar{x} = \frac{x}{r_c}, \quad \bar{r}_i = \frac{r_i}{r_c}, \quad \bar{T}_i = \frac{T_w - T_i}{T_w - T_v},$$

$$\bar{P}_{li} = \frac{P_{li}}{P_v}, \quad \bar{m} = \frac{h_{fg}}{r_c k_l (T_w - T_v)} \dot{m}, \quad \bar{\sigma} = \frac{\sigma}{\sigma_w(T_w)}.$$

Combining and rearranging the previous equations produce the following set of dimensionless coupled nonlinear ordinary differential equations,

$$\frac{d\bar{r}_i}{d\bar{x}} - D_1 = 0,$$ (3.19)

$$G_2 \frac{d\bar{\sigma}}{d\bar{T}_i} \frac{d\bar{T}_i}{d\bar{x}} - \pi_1 \frac{d\bar{P}_{li}}{d\bar{x}} + \pi_2 G_1 \bar{m} = 0,$$ (3.20)

$$\frac{d\bar{m}}{d\bar{x}} + 2\pi \bar{r}_i D_1 \frac{d\bar{T}_i}{d\bar{x}} + \frac{2\pi \bar{r}_i}{1 - \bar{r}_i} \bar{T}_i = 0,$$ (3.21)

$$\left(2\bar{r}_i^2 G_5 - G_3\right) \frac{d\bar{T}_{li}}{d\bar{x}} \frac{d\bar{P}_{li}}{d\bar{x}} + \left(G_4 - 2\bar{r}_i^2 G_6\right) D_1 \frac{d\bar{P}_{li}}{d\bar{x}}$$
$$+ \pi_3 \bar{T}_i - \bar{r}_i G_5 \frac{d\bar{\sigma}}{d\bar{T}_i} \left(\frac{d\bar{T}_i}{d\bar{x}}\right)^2 + \bar{r}_i^2 G_6 \frac{d\bar{\sigma}}{d\bar{T}_i} D_1 \frac{d\bar{T}_i}{d\bar{x}} = 0,$$ (3.22)

$$\frac{dD_1}{d\bar{x}} - \pi_4 G_7 (1 + D_1^2)^{3/2} \frac{\cos\theta}{\bar{\sigma}\bar{r}_i} \frac{d\bar{m}}{d\bar{x}} - \frac{1 + D_1}{\bar{r}_i}$$
$$- (1 + D_1^2)^{3/2} \left\{ 4\pi_1 (\bar{P}_{li} - 1) + \frac{\pi_5}{(1 - \bar{r}_i)^3} \right\} = 0,$$ (3.23)

where equation (3.19) results from reducing the order of equation (3.23). The dimensionless functions $G_1 - G_7$ are given in the Appendix A. Note that the dependent variables in this equation set are $\bar{r}_i$, $\bar{T}_i$, $\bar{m}$, $\bar{P}_{li}$, and $D_1 = d\bar{r}_i/d\bar{x}$. The dimensionless initial conditions are specified at the interline ($\bar{x} = 0$):

$$\bar{r}_{i,0} = 1 - \left(\frac{-\bar{A} T_v}{\rho_l r_c^3 h_{fg}(T_w - T_v)} \right)^{1/3} ,$$

$$\bar{T}_{i,0} = 0 ,$$

$$\bar{m}_0 = 0 ,$$

$$\bar{P}_{li,0} = 1 - \frac{1}{4\pi_1 \bar{r}_{i,0}} - \frac{\rho_l h_{fg}(T_w - T_v)}{P_v T_v} ,$$

$$D_{l,0} = 0 .$$

The above initial condition for $\bar{P}_{li,0}$ was obtained by evaluating equation (3.23) at $x = 0$, assuming the initial mean curvature was equal to $1/r_c$. This differs from most other theoretical studies which assume the mean curvature is zero at the interline, even for a capillary tube geometry.

The above expression for $r_{i,0}$ is a modification of that derived by Wayner et al. [1976] for superheated adsorbed layers. One peculiarity in this expression is that the dimensionless adsorbed layer thickness $(1 - \bar{r}_{i,0})$ goes to infinity as the superheat, $T_w - T_v$, goes to zero under isothermal conditions. Adamson and Zebib [1980] showed that adsorbed layers coupled to a meniscus should be on order of 1000 angstroms for an isothermal system with a hydrostatic head pressure, suggesting that the expression derived by Wayner et al. is accurate only to leading order. A more rigorous derivation of the superheated adsorbed layer thickness using dispersion forces can likely be obtained by applying the theory of physical adsorption pioneered by Hill [1949, 1949, 1952] in the late 1940s. Instead of taking the solution thermodynamic point of view, this approach treats the adsorbed layer and solid substrate as a single phase in equilibrium with a vapor phase.

The dimensionless groupings in equations (3.19-3.23) are:

$$\pi_1 = \frac{r_c P_v}{4\sigma_w},$$

$$\pi_2 = \frac{2\mu_l k_l (T_w - T_v)}{\pi r_c \rho_l \sigma_w h_{fg}},$$

$$\pi_3 = \frac{\alpha \mu_l}{r_c \sigma_w},$$

$$\pi_4 = \left(\frac{2-C}{2C}\right)\frac{R^{1/2} k_l (T_w - T_v)^{3/2}}{(2\pi)^{1/2}\sigma_w h_{fg}},$$

$$\pi_5 = \frac{-\overline{A}}{\sigma_w r_c^2}.$$

π_2 is the dimensionless superheat, π_3 is the Crispation number, π_4 represents the nonequilibrium charateristics of the interface, and π_5 is a dispersion number which denotes the magnitude of the dispersion forces in the thin film. The Marangoni effects in equations (3.19-3.23) are captured by terms containing $d\overline{\sigma}/d\overline{T}_i$.

3.3. Solutiⁱn Method

The numerical technique used to solve equations (3.19-3.23) incorporated both partial linearization as well as first-order backward finite differencing. The nonlinear terms in equation (3.22) were first linearized with respect to the first order derivatives of the dependent variables using a Taylor series expansion around the previous iteration. The resulting equation set was then written in matrix form with the solution vector composed of the current iteration first order derivatives. The elements in the coefficient matrix (5x5) contained constants, dependent variables, and/or previous iteration first derivatives. Similarly, the elements in the constant vector were composed of constants and/or dependent variables. A solution to the

matrix equation was obtained at each axial position by
solving the linear system, updating the dependent variables
with first-order implicit backward finite differencing, and
then iterating until the derivative terms and dependent
variables converged to a relative error of less than 10^{-8}.
In most cases, the numerical scheme converged within thirty
iterations. In order to prevent singularities in the matrix
equation and initiate the numerical scheme, the initial
values of $\overline{T}_{i,0}$, $\overline{m}_0$, and $D_{1,0}$ were set equal to 1×10^{-11}, -1
$\times 10^{-11}$, and -1×10^{-11}, respectively. Magnitudes less than
this produced physically unrealistic solutions, i.e. a
decreasing thin film thickness. An IBM PS-2 386 computer
with a mathematical coprocessor was used to generate the
numerical solutions. Run times for a complete solution were
on the order of 5 minutes for a step size of 10^{-4}. The
solution also converged at the theoretical rate when the
step size was decreased.

Solutions using equations (3.19-3.23) in the above matrix
formulation could be attained only for values of $\overline{r}_i > 0.65$,
starting from the capillary wall ($\overline{r}_{i,0} \approx 1$). This was due to
the fact that the matrix equation became stiff at this
point, forcing $dD_1/d\overline{x}$ to go to negative infinity
prematurely. For the most part, this problem was alleviated
by transforming the terms in the mean curvature as follows:

$$\overline{K} = \frac{1}{2} \left\{ \frac{1}{\overline{r}_i [1 + D_1^2]^{1/2}} - \frac{dD_1/d\overline{x}}{[1 + D_1^2]^{3/2}} \right\} = \frac{1}{2\overline{r}_i D_1} \frac{d(w\overline{r}_i)}{d\overline{x}} = \frac{1}{2\overline{r}_i} \frac{d(w\overline{r}_i)}{d\overline{r}_i}. \quad (3.24)$$

where

$$w = (1 + D_1^2)^{-1/2}.$$

Applying this transformation simply required modifying

equation (3.23) and changing one initial condition: at $\bar{x} = 0$, $D_1 = 0$, giving $(w\bar{r}_i)_0 = \bar{r}_{i,0}$. Unfortunately, a solution to the transformed equation set could not be obtained near the interline ($\bar{x} \approx 0$) because the values of D_1, which were on the order of -1×10^{-11}, were too small to produce any change in $w\bar{r}_i$ using either double-precision or extended-precision arithmetic. Quadratic-precision arithmetic was not available to us on this machine.

The problems of both initiating the algorithm and obtaining solutions near the axial centerline were eliminated by starting the solution using the original equations and then at some point changing the equations to the transformed equation set to complete the solution. The value of D_1 at which this transition occurred was chosen as 10^{-3}. Tests of the hybrid algorithm showed that order of magnitude changes on either side of this transition value produced negligible changes in the overall solution. Using the hybrid algorithm, solutions could be attained for values of $\bar{r}_i < 0.20$. A listing of the computer program which implements the hybrid algorithm is given in Appendix B.

One aspect of the physical problem that has not been discussed yet is the requirement that $D_1 \rightarrow -\infty$ as $\bar{r}_i \rightarrow 0$ (this represents the targeted endpoint for the above shooting method). Arbitrarily selecting combinations of parameters forced the far field meniscus to either curve abruptly and not intersect the axial centerline at all or form a cusp at the axial centerline. The various parameter combinations also produced different values for the mass transfer rate. Thus ensuring centerline condition is satisfied, which to our knowledge has not been done in any past studies of the evaporating meniscus in a capillary tube, is mandatory since the total mass transfer rate is strongly dependent on the system parameters.

Further tests of the hybrid algorithm showed that the

mean curvature outside the thin film was rapidly asymptotic to a constant value whose magnitude depended on the values of the parameters. This asymptotic property allowed us to define the various combinations of r_c, ΔT, and $\overline{A}$ which ensured that $D_1 \rightarrow -\infty$ as $\overline{r}_i \rightarrow 0$. When the mean curvature is constant, equation (3.24) can be integrated from some initial point on the meniscus to an arbitrary location giving

$$w\overline{r}_i - (w\overline{r}_i)_{\overline{K}=const.} = \overline{K}\,\overline{r}_i^2 - \left(\overline{K}\,\overline{r}_i^2\right)_{\overline{K}=const.} .$$

Because this expression applies at $\overline{r}_i = 0$,

$$(\overline{r}_i)_{\overline{K}=const.} = \sqrt{\left(\frac{w\overline{r}_i}{\overline{K}}\right)_{\overline{K}=const.}} . \tag{3.25}$$

Solutions which satisfy this condition will guarantee that the targeted condition of $D_1 \rightarrow -\infty$ as $\overline{r}_i \rightarrow 0$ is achieved. It is interesting to note that, in all cases, meeting this criterion require that the dimensionless mean curvature be asymptotic to one, a value typically characterizing a hemispherical meniscus. This asymptotic behavior suggests that there is likely some mathematical justification for $\overline{K} \rightarrow 1$ outside the thin film region although the proof will not be attempted here.

3.4. Results/Discussion

As stated earlier, the primary objectives of this study were to determine how the various forces at the interface influence the meniscus profile, the mass transfer rate, and the capillary pressure. As an example, room temperature hexane was chosen as the nonpolar working fluid and the accommodation coefficient was assumed to be equal to one.

The tube radius, superheat, and dispersion coefficient were chosen as independent parameters. Fluid properties were taken from Vargaftik [1975] and the range of values for the dispersion coefficient corresponded to those found in the literature (Wayner [1978], Israelachvili [1985]). Various combinations of these parameters satisfying the axial boundary condition produced dimensionless parameter ranges of $2.47 \times 10^{-11} < \pi_2 < 1.34 \times 10^{-10}$ and $2.0 \times 10^{-10} < \pi_5 < 1.0 \times 10^{-9}$. Recall, π_2 represents a dimensionless superheat while π_5 is the dispersion number. Note that all other dimensionless parameters are dependent on the material properties, the capillary radius, and/or the superheat; thus holding π_2 constant fixes all of the other dimensionless parameters except for π_5.

Figure 3.2 shows meniscus profiles for the extreme values of the dispersion number; all profiles for the intermediate values of the dispersion number fell between the two curves shown. One unexpected result shown in Figure 3.2 is that varying the dimensionless superheat, for a constant dispersion number, had no apparent effect on the meniscus profile. However, varying the dispersion number did produce a noticeable change in the meniscus profile, but only at the microscopic level near the tube wall as shown more explicitly by Figure 3.3. This is highlighted by the fact that, under all conditions, the local contact angles become equal beyond ten adsorbed layer thicknesses ($10[1 - \bar{r}_{i,0}]$). This location also loosely defines the point at which the dimensionless meniscus curvature is asymptotic to one. One noticeable difference in the microscopic characteristics of the menisci is that for larger dispersion numbers, the thin film extends further down the tube due to the stronger attractive forces between the liquid and the solid substrate. This has a significant effect on the liquid pumping capability of the capillary tube.

Figure 3.4 depicts the dimensionless mean curvature as a function of axial position. As the dispersion number increases, the dimensionless mean curvature remains at approximately 0.5 further down the tube. The first noticeable change in $\overline{K}$ occurs at $\overline{x} \approx 0.11$ for $\pi_5 = 2 \times 10^{-10}$ and $\overline{x} \approx 0.16$ for $\pi_5 = 1 \times 10^{-9}$. It is interesting to note in Figure 3.3 that the first recognizable change in the microscopic interface radial location occurs at $\overline{x} \approx 0.15$ for $\pi_5 = 2 \times 10^{-10}$ and $\overline{x} \approx 0.22$ for $\pi_5 = 1 \times 10^{-9}$, corresponding to dimensionless mean curvature values of 0.8 and 0.89, respectively. Thus, the mean curvature is very near its asymptotic value of one before any significant change in the radial interface location can be observed.

The impact that the dispersion forces and superheat on the local interfacial mass flux is shown in Figures 3.5 ($\pi_5 = 2 \times 10^{-10}$) and 3.6 ($\pi_5 = 1 \times 10^{-9}$). Both figures exhibit peaks in the dimensionless mass flux whose magnitudes tend to increase as the dimensionless superheat decreases. The reason for this counterintuitive behavior is because the dimensionless form of the mass flux (shown in each figure) is inversely proportional to the superheat. Dimensional values of the mass flux do indeed display the correct trends as the superheat increases; in fact, the dimensional mass flux is nearly directly proportional to the superheat.

Figures 3.5 and 3.6 also show that the location of the peak mass flux appears nearly constant with respect to the dimensionless superheat, although mild peak shifts occur in both curves. The peaks occur at approximately 0.12 and 0.18 for dispersion numbers of $\pi_5 = 2 \times 10^{-10}$ and $\pi_5 = 2 \times 10^{-10}$. At these locations and dispersion numbers, Figure 4 shows that the mean curvature lies in the range $0.5 < \overline{K} < 0.55$; thus, most of the evaporation occurs in the thin film region. As expected, larger dispersion numbers, which reflect the fluids affinity for the wall, produce larger

local mass fluxes or tube pumping capability.

Increasing the dispersion number produces a similar trend in the dimensionless total mass transfer rate. Figure 3.7 depicts the mass transfer rate as a function of the dispersion number for various values of the dimensionless superheat. The counterintuitive decrease in the dimensionless total mass transfer rate with increasing superheat occurs for same reasons discussed earlier. The figure shows that increasing the dispersion number dramatically increases the mass transfer rate. As mentioned earlier, the cause for the increase in mass transfer rate is due to the stronger dispersion forces between the liquid and solid substrate which increase both the thin film surface area as well as the liquid pressure difference. These results agree with previous studies which also show that the dispersion forces in the thin film significantly influence the liquid pumping capability of capillary tubes.

Although experimental verification is still required, these findings as well as others (i.e those of Wayner [1978, 1979]), suggest that the presence of surface coatings in capillary structures can either enhance or degrade capillary pumping. For example, simply adding a surface coating which increases the dispersion number over that for the uncoated solid substrate should improve capillary pumping. Similarly, preventing an oxide coating from developing on a metallic capillary structure during heat pump fabrication or due to material/fluid incompatibility should also enhance capillary pumping.

Another interesting result produced by the model is shown in Figure 3.8 which gives the dimensionless capillary pressure difference, $\overline{P}_{ui} - \overline{P}_{li}$, for all values of π_2 and π_5 as a function of capillary radius. The figure shows that the capillary pressure difference is only a function of the capillary radius, and is independent of the dispersion

number and dimensionless superheat. This observation bears
sharp contrast to the theoretical model of Wayner [1979]
which predicts that increasing the interline heat flux
reduces the capillary pressure by increasing the apparent
contact angle. In our model, neither an increase in
apparent contact angle nor a decrease in capillary pressure
was observed as the dimensionless superheat was increased.
We believe the difference between our results and those of
Wayner [1979] is the requirement that $D_1 \rightarrow -\infty$ as $\bar{r}_i \rightarrow 0$. In
fact, the capillary pressure generated by our model was
found to be constant locally and equal to $2\sigma/r_c$ up to four
decimal places. The local invariance in the capillary
pressure and the fact that the dimensionless mean curvature
is approximately equal to approximately 0.5 in the thin film
leads us to believe that the term accounting for dispersion
forces in equation (3.3) compensates for the lower curvature
in the thin film. The reason that no significant changes in
the capillary pressure in the thin film region were observed
as the dispersion coefficient was increased is due to the
initial adsorbed layer thickness: the initial adsorbed layer
thickness increases as the dispersion coefficient to the 1/3
power. This renders the dispersion term in equation (3.3)
constant at the initial condition and produces a constant
capillary pressure throughout the thin film region
regardless of the magnitude of the dispersion number.

3.5. Conclusions

A mathematical model describing the evaporating meniscus
in a capillary tube has been formulated incorporating the
full three-dimensional Young-Laplace equation, Marangoni
convection, van der Waals dispersion forces, and
nonequilibrium interface conditions. The governing
equations and boundary conditions were cast in terms of five
coupled nonlinear ordinary differential equations and solved

numerically. Material properties characterizing hexane at room temperature and pressure were used to test the model for various values of the dimensionless superheat and dispersion number.

The results showed that varying the dimensionless superheat, for a constant dispersion number (number characterizing the London-van der Waals forces in the thin film region), had no apparent effect on the meniscus profile. However, varying the dispersion number did produce a noticeable change in the meniscus profile, but only at the microscopic level near the tube wall. No change in the apparent contact angle was observed with changes in the dimensionless superheat or dispersion number. In all cases, the dimensionless mean curvature was asymptotic to a value equal to that for a hemispherical meniscus. At the microscopic scale, larger dispersion numbers produced thin films which extended further down the tube increasing the mass transfer surface area. The liquid pressure difference also increased with increasing dispersion number. These combined effects produced increases in the local interfacial mass flux and total mass transfer rate as the dispersion number was increased. Peaks in the mass flux were also observed in the thin film region. This has a significant effect on the liquid pumping capability of the capillary tube and suggests that surface coatings can play a important role in improving or degrading capillary pumping. The model also predicted that the local capillary pressure remains constant and equal to $2\sigma/r_c$ regardless of changes in the dimensionless superheat and dispersion number.

3.6. Nomenclature

Symbols

$\overline{A}$ - dispersion coefficient

D_1 - change in interface location with respect to position

F - function

G - dimensionless function

h_{fg} - latent heat of vaporization

K - mean interfacial curvature

$\dot{m}$ - liquid mass flow between the interface and tube wall
(mass flow rate)

$\dot{m}''$ - interfacial mass flux

P_{li} - pressure on the liquid side of the meniscus

P_{vi} - pressure on the vapor side of the meniscus

r - radial coordinate

r_c - tube radius

r_i - radial location of the meniscus

R - ideal gas constant divided by the molecular weight

T - temperature

T_i - interfacial temperature

u_l - liquid axial velocity

w_{li} - liquid interfacial velocity perpendicular to the
meniscus

w_{vi} - vapor interfacial velocity perpendicular to the
meniscus

x - axial coordinate

Greek

α - thermal diffusivity

μ - viscosity

ρ - density

σ - surface tension

θ - local contact angle

Subscripts

c - capillary

i – interface

l – liquid

υ – vapor

w – capillary wall

4. VAN DER WAALS FORCES IN CAPILLARY TUBES

4.1 Introduction

4.1.1 Motivation for the Study

Van der Waals forces in capillary tubes have previously
been calculated by Philip [1977b]. His study was based on
the Hamaker theory, which assumes that intermolecular forces
are pairwise additive. This approach neglects the
retardation of the van der Waals interaction, which is the
attenuation of the interaction due to the finite time
required for the interaction to propagate between atoms.
When interacting atoms are seperated by large distances,
this effect can be very significant. Also, the temperature
dependence of these forces and the presence of many-body
forces, which are important in condensed media, are not
taken into account by the Hamaker theory. Calculations using
on the Hamaker theory are often based on an unrealistic
model of the dielectric properties of the media which
assumes that photons exchanged by interacting atoms are all
of a single frequency. Electromagnetic radiation of a number
of frequencies does, in fact, contribute to the van der
Waals interaction. These difficulties suggest that there is
need for a more refined theory of van der Waals forces in
capillary tubes.

The primary objective of my research effort was to
develop a theoretical expression for the van der Waals
interaction between a condensed liquid phase and a smooth
capillary tube in the presence of a vapor-liquid interface.
The theory must account for the effect of capillary
curvature, the impact of material properties on the
interaction, retardation of the van der Waals interaction,
and temperature dependence. In many respects, then, this

work is aimed at refining the results obtained by Philip
[1977b] for the van der Waals interaction in capillary
tubes. The refinements of the present work should make it
possible to determine when the effect of capillary curvature
is comparable to the effect of surface roughness and surface
impurities on the van der Waals interaction.

4.1.2 Historical Background

In this section, the development of the theory of van
der Waals forces between macroscopic bodies is briefly
reviewed. More thorough details may be found in the reviews
and original papers cited below.

Many reviews of the theory of van der Waals forces
exist. The texts by Langbein [1974] and Mahanty and Ninham
[1976] are especially useful introductions to the subject.
Although the approach taken by Langbein is somewhat
unconventional, his text presents the theory of macroscopic
van der Waals forces in a physically clear manner. Mahanty
and Ninham's text is the single most complete reference on
van der Waals forces between macroscopic bodies. Their book
concisely summarizes the state of the art in 1976, which
includes all of the theory required for this work. The short
review article by Richmond [1975] gives a concise overview
of the field. Israelachvili [1985] has written an
exceptionally clear account of these forces. His text gives
a broad overview of surface forces, but is somewhat lacking
in details of the theory. The recent review by Barash and
Ginzburg [1989] provides both an overview of the state of
the art and an excellent list of references on the subject.
The recent advances reviewed have little bearing on the
present work, but will be of interest to the theoretician.
Finally, Nir and Vassilieff [1988] have given a concise
overview of van der Waals interactions in thin liquid films.
Their work is probably the single best reference for the

purpose of this work.

Van der Waals forces are long range intermolecular forces which arise from polarization interactions between atoms or molecules. There are three basic interactions which contribute to the total van der Waals force between bodies. These are the Keesom, the Debye, and the dispersion interactions (Israelachvili [1985]). The Keesom interaction results from the presence of interacting permanent dipoles in a medium. The Debye interaction arises when a polar molecule polarizes a nonpolar molecule, causing a dipole-induced dipole interaction. The dispersion interaction was first treated by London in 1930. It is always present, in contrast to the Keesom and Debye interactions which require the presence of permanent dipoles. The dispersion interaction results from the interaction of the instantaneous dipole moments of atoms. These are non-zero even if the time-averaged dipole moments of the atoms are zero. This interaction is quantum mechanical in origin and makes a finite time averaged contribution to the total van der Waals interaction.

At the same time that London proposed his theory of the dispersion interaction, he suggested that the van der Waals interaction given by the sum of the Debye, Keesom, and dispersion interactions could be applied to the interaction of macroscopic bodies. This idea was taken up by Hamaker [1937], who first calculated the van der Waals interaction between macroscopic bodies. He assumed that the atomic interactions are pairwise additive. The interaction energy between two macroscopic bodies can then be written in terms of a volume integral over each of the bodies:

$$E = -\int_{V_1} dV_1 \int_{V_2} dV_2 \frac{C_{12}\rho_1\rho_2}{R^6},$$

(4.1.1)

where C_{12} is a constant which characterizes the strength of the interaction. The constants ρ_1 and ρ_2 are the number densities of the types of atoms in the respective bodies.

Verwey and Overbeek [1948] developed a theory of the interaction between colloidal particles during the 1940's, which utilized the Hamaker theory to calculate an attractive term and the theory of the electric double layer to calculate the repulsive interaction. A discrepancy was found between their theoretical predictions and experimental results. Verwey and Overbeek suggested on this basis that the attractive van der Waals force between particles must decrease more rapidly than predicted by the London theory for large particle separations.

This led Casimir and Polder [1948] to explore the nature of retardation in van der Waals forces. This goal was achieved by applying quantum electrodynamics to the interaction of two neutral atoms. Casimir and Polder obtained an attractive interaction energy which varies as $1/R^7$ in the limit of large atomic separations and varies as $1/R^6$, the form predicted by London, for small separations.

The work of Casimir and Polder was the first study to properly treat the retardation of van der Waals forces. Their work also introduced an approach to calculating van der Waals forces between macroscopic bodies which is the basis for methods currently in use. The method was originally developed by Casimir and Polder to calculate the van der Waals interaction between two perfectly conducting plates separated by vacuum. Casimir and Polder began by calculating the macroscopic electromagnetic modes between the plates corresponding to thermal radiation and quantizing these modes. A divergent series was obtained for the interaction energy between the plates at a finite separation. A similar series was obtained for the limit of infinite plate separation. It was shown that the difference

of these two series, which is the change in the zero point energy of electromagnetic modes between the plates, gives a non-divergent interaction energy. This approach is basic to all modern treatments of van der Waals forces between macroscopic bodies.

Lifshitz [1955] developed a theory of the van der Waals forces between solid macroscopic media based on the fluctuation-dissipation theorem, which is described in Landau and Lifshitz [1980], pp. 386-93. In the Lifshitz theory, the van der Waals forces between media are calculated from the Maxwell stress tensor of a fluctuating electromagnetic field which is usually associated with thermal radiation. At absolute zero this field arises from zero point oscillations of the electromagnetic field (vacuum fluctuations). For the purposes of this theory, the properties of the media are completely specified by their dielectric permittivities, $\epsilon_i(\omega)$, where the subscripts i denote the particular media. The Lifshitz theory represents a major advance over previous efforts because it accounts for the temperature dependence and retardation of van der Waals forces in a natural way. The Lifshitz theory also accounts for the contribution of many-body forces to the van der Waals forces, which are quite significant in condensed media. This provides a major advantage over the approach of Hamaker which is only applicable to rarified media.

The Lifshitz theory was generalized by Dzyaloshinskii et al [1961] through application of quantum field theoretic methods to calculate the effect of long wavelength electromagnetic fluctuations on a general condensed medium. A detailed account of this approach has been given in the text by Abrikosova et al [1963]. The theory of Dzyaloshinskii et al represents a powerful extension of the Lifshitz theory because it is applicable to any medium, including thin liquid films (Dzyaloshinskii et al [1959]).

In fact, predictions of the thickness of liquid helium films were shown to compare quite favorably with known experimental val s in the 1961 paper. Practical application of the theory, h ever, entails the use of a mathematically cumbersome Green's function formalism. For this reason, the theory never attained widespread use.

The practical difficulties of the Lifshitz theory were first overcome by van Kampen et al [1968] who suggested that the van der Waals forces between two semi-infinite media could be calculated from the macroscopic electromagnetic surface modes of the system. The interaction between the media was expressed as the difference between the sum of the zero point oscillations of all modes when the media have a finite separation and the sum of zero point oscillations of all modes when the media have an infinite separation. Clearly the van Kampen surface mode analysis approach is an extension of the aforementioned approach of Casimir and Polder.

Ninham, Parsegian, and Weiss [1970] extended the analysis of van Kampen et al to account for temperature-dependence and retardation. Retardation was incorporated by solving Maxwell's equations for the electromagnetic surface modes rather than the Laplace equation, which corresponds to an electrostatic problem. This approach was shown by Ninham and Parsegian [1970b] to be equivalent to that of Dzyaloshinskii et al [1961]. The surface mode analysis approach merely requires solution of Maxwell's equations for electromagnetic surface modes, from which a dispersion equation is developed. The dispersion equation, which relates the frequencies of the surface modes to their corresponding wavevectors, can then be used to directly calculate the free energy of interaction. A state density integral expression developed by Ninham, Parsegian, and Weiss [1970] is used to obtain the free energy of

interaction. This approach stands in sharp contrast to the cumbersome field theoretic formulation of Dzyaloshinskii et al.

During the early 1970's a great deal of work was done which further validated the work of Ninham, Parsegian, and Weiss. Perhaps the most important effort in this regard was the work of Gerlach [1971], who showed that the van der Waals force between two solid bodies can be calculated from the change in zero point energies of interacting surface electrons undergoing plasma oscillations about their ion cores. Gerlach's paper elaborates on the fundamental mechanism of the interaction, as well as illustrating how the surface mode analysis method can be derived from quantum mechanical many-body theory. This theme was pursued by a number of researchers in the 1970's, whose efforts have been summarized by Barton [1979].

Progress in understanding van der Waals forces has by no means been limited to theory. The surface mode analysis approach was applied to calculating the van der Waals forces across thin films of liquid helium by Richmond and Ninham [1976]. Their predictions are well in agreement with the measurements of Sabisky and Anderson [1973]. Predictions of the disjoining pressure of hydrocarbon films adsorbed on water made by Richmond, Ninham, and Ottewill [1973] were also shown to be in good agreement with experimental results by Mahanty and Ninham [1976].

Perhaps the most compelling evidence was produced by Israelachvili and Tabor [1972], who directly measured the van der Waals forces between crossed mica cylinders. Crossed cylinders were used so that the surfaces could be positioned with precision by use of a piezoelectric crystal to which the cylinders were attached. The seperation of the cylinders could then be measured by optical diffraction. The surface force between the cylinders was measured using a sensitive

spring. This method of directly measuring the van der Waals forces between surfaces continues to be used. More recently, Horn and Israelachvili [1981] have used a similiar apparatus to measure the surface forces between mica cylinders seperated by an electrolyte solution. They were clearly able to resolve not only the van der Waals forces, but electric double-layer and solvation forces as well. These measurements were carried out down to seperations of less than 1.0 nm with a resolution of 0.1 nm. A detailed description of the experimental method is given in the original paper by Horn and Israelachvili.

4.2 Theory

4.2.1 Assumptions of the Theory

The van der Waals forces in a smooth capillary tube have previously been studied by Philip [1977b] using the Hamaker theory. Because the condensed liquid phase in a capillary tube can be relatively thick (several hundred angstroms), retardation effects have a significant impact on the van der Waals interaction at the vapor-liquid interface in a capillary tube. The temperature dependence of the van der Waals forces is also an important factor in assessing their role in capillary pumping for heat pipe wick applications. These factors make the Hamaker theory a poor approximation for this problem. A more refined approach based on the van Kampen surface mode analysis method will be developed in this chapter.

The theory developed in this chapter is based on three key assumptions. These assumptions may be summarized as follows: (1) the condensed liquid phase thickness is taken to be greater than 10 nm; (2) the electric double layer forces are negligible at the vapor-liquid interface; and (3)

the surface of the capillary tube is treated as an ideal surface. In this context an ideal surface is a smooth, defect free surface devoid of chemical impurities.

The first assumption has been made on the basis of preliminary calculations of the thickness of condensed liquid phases, which should be in the range of 20-100 nm. This justifies omission of short range surface forces. The first assumption also justifies the use of the surface mode analysis method for this problem since a sufficiently thick condensed liquid phase can be treated in a continuum approximation.

Omission of electric double-layer forces and solvation forces is reasonable for thick liquid films. Electric double layer forces arise from the interaction of counter-ions in the condensed liquid phase with charge on a solid surface. These forces tend to decay rapidly away from the surface because source charges remain local to the surface and screening takes place. For a condensed liquid phase having a thickness on the order of 10 nm, it may be necessary to correct the present theory for double-layer forces, if the charge density of the capillary tube surface is sufficiently large.

Idealization of the capillary tube surface is motivated by mathematical convenience rather than a belief that such a model is strictly true. Certainly the effect of surface defects on the van der Waals interaction will be small for a thick condensed liquid phase. By calculating the effect of capillary curvature on the van der Waals forces, an estimate of when these forces are comparable to surface roughness effects may be arrived at. The effect of surface roughness on the van der Waals interaction was initially studied by van Bree [1974]. The work of Maradudin and Mazur [1980] indicates that on a rough surface the magnitude of the van

der Waals forces is dependent on the ratio of the root-mean-square asperity height of the surface to the transverse correlation length of the surface (the distance between consecutive peaks and valleys on the surface) and the ratio of the transverse correlation length to the thickness of the condensed liquid phase. For a typical surface, Maradudin and Mazur give a ratio of transverse correlation length to root-mean-square asperity height of approximately 0.1. The deviation of the van der Waals free energy per unit area from that of a flat plate is approximately 2% when the thickness of the condensed liquid phase is six times the transverse correlation length of the surface. These values, which correspond to a relatively thin condensed liquid phase, illustrate that surface roughness effects a small correction to results obtained for smooth surfaces so long as the thickness of the condensed liquid phase is greater than 10 nm or so.

4.2.2 Surface Modes in a Capillary Tube

Application of the van Kampen surface mode analysis method requires a dispersion equation for surface electromagnetic modes. The usefulness of the dispersion equation lies in the fact that its zeros correspond to the normal modes of the system, from which one can readily find the interaction energy. A dispersion equation can be derived for a given system by solving Maxwell's equations for the particular geometry, then matching potentials and fields across the boundaries between media. A system of equations is then obtained from which the dispersion equation is found.

An equivalent approach to the one just described is to solve the electromagnetic wave equation for the vector and scalar potentials corresponding to surface modes. This is

the approach that is commonly used in the literature. Working in the Lorentz gauge and assuming no source charges or currents, which is consistant with exclusion of electric double-layer effects at the vapor-liquid interface, one must then solve

$$\nabla^2 \vec{A} - \mu\sigma\frac{\partial \vec{A}}{\partial t} - \mu\epsilon\frac{\partial^2 \vec{A}}{\partial t^2} = 0,$$

(4.2.1a)

$$\nabla^2 \phi - \mu\sigma\frac{\partial \phi}{\partial t} - \mu\epsilon\frac{\partial^2 \phi}{\partial t^2} = 0,$$

(4.2.1b)

$$\nabla \cdot \vec{A} - \mu\sigma\phi - \mu\epsilon\frac{\partial \phi}{\partial t} = 0,$$

(4.2.1c)

in each of the media. The required dispersion equations for the surface electromagnetic modes are then obtained by satisfying the boundary conditions between the media.

For the present study, eqns. (4.2.1) must be solved for a geometry of two concentric cylinders corresponding to the vapor phase and the condensed liquid phase surrounded by a third medium which represents to capillary tube material. The capillary tube medium may be considered to extend out to infinity so long as the thickness of the capillary tube is greater than the skin depth of the electromagnetic field. Otherwise the finite thickness of the capillary tube will affect the electromagnetic surface modes at the interface between the capillary tube and the condensed liquid phase within the tube.

A convenient method of solving eqns. (4.2.1) for surface modes has been given by Richmond and Ninham [1971]. Their approach entails three steps. First, eqns. (4.2.1) are Fourier transformed with respect to time which yields a system of the form

$$\nabla^2 \vec{A}_\omega + \mu\epsilon\omega^2\left(1 + \frac{i\sigma}{\epsilon\omega}\right)\vec{A}_\omega = 0,$$

(4.2.2a)

$$\nabla^2 \phi_\omega + \mu\epsilon\omega^2\left(1 + \frac{i\sigma}{\epsilon\omega}\right)\phi_\omega = 0, \tag{4.2.2b}$$

$$\nabla \cdot \vec{A}_\omega - i\mu\epsilon\omega\left(1 + \frac{i\sigma}{\epsilon\omega}\right)\phi_\omega = 0, \tag{4.2.2c}$$

Eqns. (4.2.2) may then be simplified by the introduction of a 'superpotential' (also known as a "Hertz vector" - see Panofsky and Phillips [1962], pp. 254-5) with Fourier components $\vec{Z}_\omega$, defined by the relations

$$\phi_\omega = -\nabla \cdot \vec{Z}_\omega \quad , \quad \vec{A}_\omega = -i\mu\epsilon\omega\left(1 + \frac{i\sigma}{\epsilon\omega}\right)\vec{Z}_\omega. \tag{4.2.3}$$

Richmond and Ninham have shown how this leads to a substantial simplification of the problem. In particular, the gauge condition, eqn. (4.2.2c) reduces to an identity and the field equations, eqns. (4.2.2a) and (4.2.2b) may be combined into a vector Helmholtz equation of the form

$$\nabla^2 \vec{Z}_\omega + \mu\epsilon\omega^2\left(1 + \frac{i\sigma}{\epsilon\omega}\right)\vec{Z}_\omega = 0, \tag{2.4}$$

A useful discussion of the vector Helmholtz equation in cylindrical coordinates may be found in Moon and Spencer [1938], pp. 138-9. The third step in the approach of Richmond and Ninham is to introduce a superpotential corresponding to surface modes into eqn. (4.2.4) and solve the resulting equation. In cylindrical coordinates surface modes are defined by

$$\vec{Z}_\omega = e^{i(kz+m\theta)}\vec{Z}_{surf}(r), \tag{4.2.5}$$

Substitution of eqn. (4.2.5) into eqn. (4.2.4) yields a system of ordinary differential equations of the form

$$\frac{d^2 Z_r}{dr^2} + \frac{1}{r}\frac{dZ_r}{dr} - \left(\frac{m^2+1}{r^2} + \rho_i^2\right)Z_r = 0, \tag{4.2.6a}$$

$$\frac{d^2 Z_\theta}{dr^2} + \frac{1}{r}\frac{dZ_\theta}{dr} - \left(\frac{m^2-}{r^2} \quad \rho_i^2\right)Z_\theta = 0, \tag{4.2.6b}$$

$$\frac{d^2 Z_z}{dr^2} + \frac{1}{r}\frac{dZ_z}{dr} - \left(\frac{m^2}{r^2} + \rho_i^2\right)Z_z = 0,$$
(4.2.6c)

where

$$\rho_i^2 = k^2 - \mu\epsilon\omega^2\left(1 + \frac{i\sigma}{\omega\epsilon}\right),$$
(4.2.7)

in each of the media. The subscript i is used to denote the medium in question. For the purposes of this work, an i value of 1 will refer to the capillary tube medium, 3 will denote the condensed liquid phase, and 2 will denote the vapor phase above the liquid. It is also important to bear in mind that the dielectric susceptibility of each of these media will be a function of frequency, $\epsilon_i = \epsilon_i(\omega)$.

The system of equations (4.2.6) may be simplified for situations of practical interest because solutions with relatively large arguments and large orders arise when the condensed liquid phase thickness is much less than the capillary radius. Under these circumstances the approximation $m^2 + 1 \approx m^2$ may be used. Eqns. (4.2.6) may then be re-written in the compact vector form

$$\frac{d^2 \vec{Z}_{surf}}{dr^2} + \frac{1}{r}\frac{d\vec{Z}_{surf}}{dr} - \left(\frac{m^2}{r^2} + \rho_i^2\right)\vec{Z}_{surf} = 0,$$
(4.2.8)

General solutions of eqn. (4.2.8) have the form

$$Z_{surf} = a_m I_m(\rho_2 r) + b_m K_m(\rho_2 r) \quad for \quad r < r_i,$$
(4.2.9a)

$$Z_{surf} = c_m I_m(\rho_3 r) + d_m K_m(\rho_3 r) \quad for \quad r_i \le r \le r_c,$$
(4.2.9b)

$$Z_{surf} = e_m I_m(\rho_1 r) + f_m K_m(\rho_1 r) \quad for \quad r > r_c,$$
(4.2.9c)

where r_c is the radius of the capillary tube and r_i is the radius of the vapor-liquid interface within the capillary tube. These solutions must satisfy the boundary conditions that $\vec{Z}_{surf}$ is finite in the limit as r goes to zero and that it is zero in the limit as r goes to infinity. These conditions are satisfied by requiring that $\vec{b}_m = 0$ and

$\vec{e}_m = 0$ *for all m. Matching potentials and fields at* $r = r_i$ *a system is obtained having the form*

$$\epsilon_2 a_r I_m(\rho_2 r_i) = \epsilon_3 [c_r I_m(\rho_3 r_i) + d_r K_m(\rho_3 r_i)], \qquad (4.2.10a)$$

$$\epsilon_2 a_\theta I_m(\rho_2 r_i) = \epsilon_3 [c_\theta I_m(\rho_3 r_i) + d_\theta K_m(\rho_3 r_i)], \qquad (4.2.10b)$$

$$\epsilon_2 a_z I_m(\rho_2 r_i) = \epsilon_3 [c_z I_m(\rho_3 r_i) + d_z K_m(\rho_3 r_i)], \qquad (4.2.10c)$$

$$\rho_2 a_r I_m'(\rho_2 r_i) = \rho_3 [c_r I_m'(\rho_3 r_i) + d_r K_m'(\rho_3 r_i)], \qquad (4.2.10d)$$

$$\rho_2 \epsilon_2 a_\theta I_m'(\rho_2 r_i) = \rho_3 \epsilon_3 [c_\theta I_m'(\rho_3 r_i) + d_\theta K_m'(\rho_3 r_i)], \qquad (4.2.10e)$$

$$\rho_2 \epsilon_2 a_z I_m'(\rho_2 r_i) = \rho_3 \epsilon_3 [c_z I_m'(\rho_3 r_i) + d_z K_m'(\rho_3 r_i)], \qquad (4.2.10f)$$

Note that the subscipt m has been dropped in these expressions since they must hold for all m. Also, the vector components of the coefficents in eqn. (4.2.9) explicitly appear here. In a similiar fashion, matching fields and potentials at $r = r_c$ yields a system

$$\epsilon_3 [c_r I_m(\rho_3 r_c) + d_r K_m(\rho_3 r_c)] = \epsilon_1 f_r K_m(\rho_1 r_c), \qquad (4.2.11a)$$

$$\epsilon_3 [c_\theta I_m(\rho_3 r_c) + d_\theta K_m(\rho_3 r_c)] = \epsilon_1 f_\theta K_m(\rho_1 r_c), \qquad (4.2.11b)$$

$$\epsilon_3 [c_z I_m(\rho_3 r_c) + d_z K_m(\rho_3 r_c)] = \epsilon_1 f_z K_m(\rho_1 r_c), \qquad (4.2.11c)$$

$$\rho_3 [c_r I_m'(\rho_3 r_c) + d_r K_m'(\rho_3 r_c)] = \rho_1 f_r K_m'(\rho_1 r_c) \qquad (4.2.11d)$$

$$\rho_3 \epsilon_3 [c_\theta I_m'(\rho_3 r_c) + d_\theta K_m'(\rho_3 r_c)] = \rho_1 \epsilon_1 f_\theta K_m'(\rho_1 r_c), \qquad (4.2.11e)$$

$$\rho_3 \epsilon_3 [c_z I_m'(\rho_3 r_c) + d_z K_m'(\rho_3 r_c)] = \rho_1 \epsilon_1 f_z K_m'(\rho_1 r_c), \qquad (4.2.11f)$$

The above two systems can be solved for each vector component of $\vec{a}$ and $\vec{f}$ yielding the following system of equations:

$$\left[\frac{\epsilon_3 I_m(\rho_3 r_i)}{\epsilon_2 I_m(\rho_2 r_i)} - \frac{\rho_3 I_m'(\rho_3 r_i)}{\rho_2 I_m'(\rho_2 r_i)} \right] b_r + \left[\frac{\epsilon_3 K_m(\rho_3 r_i)}{\epsilon_2 I_m(\rho_2 r_i)} - \frac{\rho_3 K_m'(\rho_3 r_i)}{\rho_2 I_m'(\rho_2 r_i)} \right] c_r = 0, \qquad (4.2.12a)$$

$$\left[\frac{\epsilon_3 I_m(\rho_3 r_c)}{\epsilon_1 K_m(\rho_1 r_c)} - \frac{\rho_3 I_m'(\rho_3 r_c)}{\rho_1 K_m'(\rho_1 r_c)} \right] b_r + \left[\frac{\epsilon_3 K_m(\rho_3 r_c)}{\epsilon_1 K_m(\rho_1 r_c)} - \frac{\rho_3 K_m'(\rho_3 r_c)}{\rho_1 K_m'(\rho_1 r_c)} \right] c_r \qquad (4.2.12b)$$

$$\left[\frac{I_m(\rho_3 r_i)}{I_m(\rho_2 r_i)} - \frac{\rho_3 I_m'(\rho_3 r_i)}{\rho_2 I_m'(\rho_2 r_i)} \right] b_\theta + \left[\frac{K_m(\rho_3 r_i)}{I_m(\rho_2 r_i)} - \frac{\rho_3 K_m'(\rho_3 r_i)}{\rho_2 I_m'(\rho_2 r_i)} \right] c_\theta = 0, \qquad (4.2.12c)$$

$$\left[\frac{I_m(\rho_3 r_c)}{K_m(\rho_1 r_c)} - \frac{\rho_3 I_m{'}(\rho_3 r_c)}{\rho_1 K_m{'}(\rho_1 r_c)}\right] b_\theta + \left[\frac{K_m(\rho_3 r_c)}{K_m(\rho_1 r_c)} - \frac{\rho_3 K_m{'}(\rho_3 r_c)}{\rho_1 K_m{'}(\rho_1 r_c)}\right] c_\theta = 0 \quad (4.2.12d)$$

Note that a system identical to eqns (4.2.12c) and (4.2.12d), but corresponding to z component has been omitted since it will give the same results as these equations. Non-trivial solutions of these systems are subject to the requirement that their secular determinants be equal to zero. The secular determinant corresponding to the radial components of these coefficients reduces to

$$D_1(\omega; k_z, m) = 1 - \overline{\Delta}_{13}\overline{\Delta}_{23}\frac{I_m(\rho_3 r_c)K_m(\rho_3 r_c)}{K_m(\rho_3 r_i)I_m(\rho_3 r_i)}. \quad (4.2.13a)$$

In a similar fashion, the two components that are tangential to the surface of the capillary tube yield identical secular determinants of the form

$$D_2(\omega; k_z, m) = 1 - \Delta_{13}\Delta_{23}\frac{I_m(\rho_3 r_i)K_m(\rho_3 r_c)}{K_m(\rho_3 r_i)I_m(\rho_3 r_c)}. \quad (4.2.13b)$$

Eqns. (4.2.13) are the dispersion equations for the problem at hand. In these equations the delta terms are defined as follows :

$$\overline{\Delta}_{13} = \frac{\rho_1 \epsilon_3 - \rho_3 \epsilon_1 \dfrac{K_m{'}(\rho_3 r_c)K_m(\rho_1 r_c)}{K_m(\rho_3 r_c)K_m{'}(\rho_1 r_c)}}{\rho_1 \epsilon_3 - \rho_3 \epsilon_1 \dfrac{I_m{'}(\rho_3 r_c)K_m(\rho_1 r_c)}{I_m(\rho_3 r_c)K_m{'}(\rho_3 r_c)}}, \quad (4.2.14a)$$

$$\overline{\Delta}_{23} = \frac{\rho_2 \epsilon_3 - \rho_3 \epsilon_2 \dfrac{I_m{'}(\rho_3 r_i)I_m(\rho_2 r_i)}{I_m(\rho_3 r_i)I_m{'}(\rho_2 r_i)}}{\rho_2 \epsilon_3 - \rho_3 \epsilon_2 \dfrac{K_m{'}(\rho_3 r_i)I_m(\rho_2 r_i)}{K_m(\rho_3 r_i)I_m{'}(\rho_2 r_i)}}, \quad (4.2.14b)$$

$$\Delta_{13} = \frac{\rho_1 - \rho_3 \dfrac{K_m{'}(\rho_3 r_c)K_m(\rho_1 r_c)}{K_m(\rho_3 r_c)K_m{'}(\rho_1 r_c)}}{\rho_1 - \rho_3 \dfrac{I_m{'}(\rho_3 r_c)K_m(\rho_1 r_c)}{I_m(\rho_3 r_c)K_m{'}(\rho_3 r_c)}}, \quad (4.2.14c)$$

$$\Delta_{23} = \frac{\rho_2 - \rho_3 \dfrac{I_m{'}(\rho_3 r_i)I_m(\rho_2 r_i)}{I_m(\rho_3 r_i)I_m{'}(\rho_2 r_i)}}{\rho_2 - \rho_3 \dfrac{K_m{'}(\rho_3 r_i)I_m(\rho_2 r_i)}{K_m(\rho_3 r_i)I_m{'}(\rho_2 r_i)}}. \quad (4.2.14d)$$

Note that the primes on the modified Bessel functions in

these equations denote derivatives with respect to r.

The dispersion equation defined by eqns. (4.2.13) and (4.2.14) can be simplified for most cases of practical interest. Because an integer number of electromagnetic surface modes must span the inner circumference of the capillary tube, it follows that $k_\theta = m/r_c$, where m denotes the m-th mode. On the other hand, Lifshitz [1955] has shown that the wavenumbers corresponding to surface modes have magnitudes of the order $1/l$, where l is the thickness of the condensed liquid phase. On the basis of these two facts, it can be concluded that the number of theta modes contributing to the van der Waals interaction goes as r_c/l. When the adsorbed layer thickness is much smaller than the capillary radius, a large number of theta modes contribute to the van der Waals interaction and the arguments of the modified Bessel functions in eqns. (4.2.13) and (4.2.14) are large. Considerable simplification of the dispersion equation can then be attained by the use of asymptotic expansions. Using the expansions for large orders given in Abramowitz and Stegan [1972], one finds that

$$I_m(\rho_3 r_c) \sim \frac{e^{m\eta}}{\sqrt{2\pi k r_c}}\left[1 + \frac{1}{24k}\left(3 - \frac{5k_\theta^2}{k^2}\right)\right], \qquad (4.2.15a)$$

$$K_m(\rho_3 r_c) \sim \sqrt{\frac{\pi}{2kr_c}}e^{-m\eta}\left[1 - \frac{1}{24kr_c}\left(3 - \frac{5k_\theta^2}{k^2}\right)\right], \qquad (4.2.15b)$$

where

$$m\eta = r_c\sqrt{k_\theta^2 + \rho_i^2} + m\ln\frac{\rho_i}{k_\theta + \sqrt{k_\theta^2 + \rho_i^2}}, \qquad (4.2.15c)$$

In these results only terms of first order in l/r_c and $1/kr_c$ have been retained. Higher terms are quite small so long as $l/r_c < 0.1$. For modified Bessel functions of argument $\rho_3 r_i$ one may use the ansatz $\rho_3 r_i = \rho_3 r_c\left(1 - \frac{l}{r_c}\right)$. Note that the quantity

$k = \sqrt{k_\theta^2 + \rho_3^2}$ has been defined above.

Using the approximations given by eqns. (4.2.15) one finds that

$$\frac{K_m(\rho_3 r_c) I_m(\rho_3 r_i)}{I_m(\rho_3 r_c) K_m(\rho_3 r_i)} \approx \left(1 + \frac{l}{r_c} k_\theta\right) e^{-2kl}, \tag{4.2.16}$$

Eqns. (4.2.13) may be re-written using eqn. (4.2.18). The results are that

$$D_1(\omega; k_z, k_\theta) = 1 - \overline{\Delta}_{13} \overline{\Delta}_{23} \left(1 + \frac{l}{r_c} k_\theta\right) e^{-2kl}, \tag{4.2.17a}$$

$$D_2(\omega; k_z, k_\theta) = 1 - \Delta_{13} \Delta_{23} \left(1 + \frac{l}{r_c} k_\theta\right) e^{-2kl}, \tag{4.2.17b}$$

Simplification of eqns. (4.2.14) follows from the observation that

$$\frac{I_m{}'(x)}{I_m(x)} = \frac{d}{dx} \ln I_m(x), \tag{4.2.18a}$$

and

$$\frac{K_m{}'(x)}{K_m(x)} = \frac{d}{dx} \ln K_m(x), \tag{4.2.18b}$$

where prime denotes derivatives with respect to x. Substituting the asymptotic expansion for modified Bessel functions of large order into these expressions, one finds that the logarithmic derivatives are well approximated by

$$\frac{I_m{}'(\rho_i r_c)}{I_m(\rho_i r_c)} = +1, \tag{4.2.19a}$$

$$\frac{K_m{}'(\rho_i r_c)}{K_m(\rho_i r_c)} = -1, \tag{4.2.19b}$$

Substitution of these last two results into eqns. (4.2.14) yields the flat plate expressions for these terms given in Appendix C. The only deviation of the dispersion equations eqns. (4.2.17) from the flat plate result is given by the

l/r_c term in these equations. The aforementioned term vanishes in the limit that l/r_c approaches zero and the flat plate dispersion equations are recovered.

4.2.3 Calculation of the Free Energy

Ninham, Parsegian, and Weiss [1970] have shown that the van der Waals free energy of interaction for a given geometry can be found directly from the corresponding dispersion equation for surface electromagnetic modes. According to their work, the interaction free energy per unit area, $F(l,T)$, is given by a state density sum of the form

$$F(l,T) = \frac{K_B T}{(2\pi)^2} \sum_{n=0}^{\infty} {}' \sum_{\vec{k}_{surf}} \ln D(i\xi_n, \vec{k}_{surf}),$$ (4.2.20)

where l corresponds to the thickness of the condensed liquid phase, T is temperature, $D(i\xi_n, \vec{k}_{surf})$ is the dispersion equation for the system under consideration, and $\vec{k}_{surf}$ are the surface modes of the system. The sum over n in eqn. (4.2.20) is taken over imaginary frequencies defined by $i\xi_n = 2\pi n K_B T/\hbar$. The prime on the first sum indicates the only half the first term contributes to the overall sum. The reader should note that eqn. (4.2.20) is simply a sum over states for an interacting oscillator system.

For the problem at hand, Eqn. (4.2.20) may be written

$$F(l,T) = \frac{K_B T}{(2\pi)^2} \sum_{n=0}^{\infty} {}' \sum_{k_\theta} \sum_{k_z} \ln D(i\xi_n; k_\theta, k_z).$$ (2.21)

At this juncture it is convenient to introduce a change of variables to non-dimensional variables of the form

$$k_z = \left(\frac{\xi_n}{\xi_s}\right)\frac{p_z}{2l} \quad , \quad k_\theta = \left(\frac{\xi_n}{\xi_s}\right)\frac{p_\theta}{2l}.$$ (4.2.22)

A procedure similiar to this is used in developing results for a flate plate geometry (Mahanty and Ninham, [1976]). Using the change of variables defined by eqn. (4.2.22) and noting that the surface modes are closely spaced so that the sums in Eqn. (4.2.21) may be replaced by integrals, an expression is generated for the interaction free energy per unit area of the form

$$F(l,T) = \frac{K_B T}{(2\pi l)^2} \sum_{n=0}^{\infty} {}' \left(\frac{\xi_n}{\xi_s}\right)^2 \int_{-\infty}^{\infty} dp_\theta \int_{-\infty}^{\infty} dp_z \ln D(i\xi_n; p_\theta, p_z). \quad (4.2.23)$$

This expression applies to condensed liquid phases in a smooth capillary tubes in the presence of a vapor-liquid interface. Note that in eqn. (4.2.23) the dispersion equation $D(i\xi_n; p_\theta, p_z)$ is that which was derived in the previous section.

It has already been demonstrated that the dispersion equation derived above for a capillary tube properly asymptotes to the flat plate limit. The Lifshitz result for interaction energy per unit area in a flat plate geometry (Appendix C) can be recovered from the theory developed here, as well. Eqn. (4.2.23) yields the flat plate result in the asymptotic limit because the integrand is only a function of $p = \sqrt{p_\theta^2 + p_k^2}$. Introducing a change of variables of the form $p'^2 = p^2 - 1$ and applying the prescription

$$\int_{-\infty}^{+\infty} dp_\theta \int_{-\infty}^{+\infty} dp_z \rightarrow 2\pi \int_1^{+\infty} dp', \quad (4.2.24)$$

the Lifshitz expression for the free energy of interaction per unit area for a flat plate geometry is recovered from eqn. (4.2.23).

4.3 Discussion

4.3.1 Models of Dielectric Behavior
Dielectric permittivities must be calculated for both

the capillary wall material, which we assume to be a metal for purposes of this work, and for both the liquid and vapor within the capillary tube. The behavior of these different media can be modelled using two different methods, the Kramers-Kronig analysis technique and the Ninham-Parsegian representation.

The Kramers-Kronig analysis method for calculating $\epsilon(i\xi)$ is based on the application of the relation

$$\epsilon(i\xi) = 1 + \frac{2}{\pi} \int_0^\infty \frac{d\omega\,\omega\,\epsilon''(\omega)}{\omega^2 + \xi^2},\qquad (4.3.1)$$

which connects the complex part of the dielectric permittivity, $\epsilon''(\omega)$, to the dielectric permittivity of a complex argument, $\epsilon(i\xi)$. A detailed derivation of eqn. (4.3.1) is given by Landau and Lifshitz [1980], pp.380-3. This relation is useful because $\epsilon''(\omega)$ can be obtained from optical reflection measurements of the index of refraction, $n(\omega)$, and extinction coefficient, $k(\omega)$, of a given material via the relationship $\epsilon''(\omega) = 2n(\omega)k(\omega)$.

Lifshitz [1955] originally suggested the use of the Kramers-Kronig method in calculating van der Waals forces between macroscopic media. Application of the Kramers-Kronig method to the calculation of van der Waals forces in thin liquid films is discussed in detail by Pashley [1977]. The first practical calculations using this method were done by Krupp, Schnabel and Walter [1972], who studied the relationship between the quality of optical data used and the accuracy of the corresponding van der Waals force calculation. Krupp, Schnabel and Walter also noted that $\epsilon''(\omega)$ must be extrapolated in the low frequency region since optical spectrum measurements do not extend to sufficiently low frequencies. A power fit of available data was used for this purpose in the present study. A number of trial calculations indicated that the results are not very

sensitive to the particular choice of extrapolation function.

Practical application of eqn. (4.3.1) is difficult because available optical data have not been measured at equi-spaced intervals. Standard numerical integration techniques cannot be applied to the evaluation of eqn. (4.3.1) as a result. This difficulty was resolved by modifying the trapezoidal rule to accommodate a variable step size. This procedure is similar to that used in other variable step size methods. Eqn. (4.3.1) must, however, be evaluated over a data set of discrete points rather than a known function. The trapezoidal rule was used because the optical spectra being integrated tend not to be smooth functions. The extrapolation function was numerically integrated up to the lowest frequency of the available optical data using a composite Simpson method, as described by Burden and Faires [1985], p. 165.

The Ninham-Parsegian representation method was originally developed by Ninham and Parsegian [1970a] for use in a study of van der Waals forces across triple-layer films. The fundamental assumption used in deriving the Ninham-Parsegian representation is that the absorption spectrum of the medium in question consists of sharp, well defined peaks and can thus be represented as a sum of Dirac delta functions. A representation for the dielectric permittivity can then be obtained by application of the Kramers-Kronig analysis (see Hough and White [1980] for details). The dielectric permittivity is expressed as a sum over absorption peaks having the form

$$\epsilon(i\xi) = \sum_i \frac{c_i}{1 + (\xi/\omega_i)^2},$$
(4.3.2)

where the ω_i are frequencies corresponding to absorption peaks and the C_i are constants which can be calculated from readily available optical and dielectric constants for the medium. In essence, this method is used to interpolate the dielectric susceptibility between known values.

The Ninham-Parsegian approach has been subject to some criticism, most notably by Nir et al [1972], who claimed that the method is only valid for sufficiently rarified media. This claim was shown to be clearly false by Gingell and Parsegian [1972], who responded to all criticisms of the time. The method was placed on a firm experimental basis by the extensive work of Hough and White [1980]. Their work, which is the most useful reference on the subject, proposes minor changes to the original prescription of Ninham and Parsegian which are justified by extensive analysis and experimental documentation.

4.3.2 The Computer Program

The theory developed in section 4.2 and the aforementioned models of dielectric properties have been implemented in a computer program which is listed in Appendix D. The program was designed to run on a personal computer with a numeric coprocessor. It was written using Turbo Pascal version 5.0 (Turbo Pascal is a registered trademark of Borland International). Development and testing of the program was carried out in a modular fashion so that individual procedures were tested before being integrated into the main program. The program has thus been proven to be reliable while also being readily expandable. The routines in the program fall into three broad categories: general purpose procedures, procedures for modelling dielectric behavior, and the procedures used which are actually used to calculate the van der Waals interaction. The physical constants used in the program are from Weast

[1981].

Three general purpose routines are used by the program. The first of these routines displays an option menu which accepts the parameters for a given program run. When the user specifies which metal the capillary tube is made of, another procedure is used to open the appropriate data file and load the optical spectrum data into an array which is used later in the program to calculate dielectric permittivities. A third general purpose procedure is used throughout the main program to display an appropriate error message and terminate program execution if an error is trapped by the program. Both user input selections and calculated results are tested throughout the program to insure that valid parameters are being used.

Dielectric properties are calculated by a single procedure in the program which in turn invokes other functions to calculate the dielectric permittivities. The Ninham-Parsegian representations for each condensed liquid phase selection specified in the menu are implemented as individual functions which are called by the aforementioned routine. Several functions are required for implementation of the Kramers-Kronig analysis method used the calculate the dielectric permittivities of the capillary tube itself. The trapezoidal rule is used by a main routine to evaluate eqn. (4.3.1), the required optical spectrum being available to the routine as an array, as was previously mentioned. An extrapolation function is used in conjunction with a composite Simpson integration routine to estimate the low frequency contribution to $\epsilon(i\xi_n)$.

A single function was implemented to calculate the natural logarithm of the dispersion equation as given by eqns. (4.2.17). This function is repeatedly called by a double-integral routine which evaluates the state-density integral in eqn. (4.2.23). An effort was made to optimize

this function, because it is called numerous times during a typical run. A 32-point Gauss-Legendre integration algorithm was developed for this purpose. This method was chosen because the integration function is smooth and decays exponentially with respect to the total surface wave number. Of equal importance, this method runs quickly. Given the number of integral evaluations for required for the problem at hand, this factor is significant. Gauss-Legendre integration is discussed in Press et al [1989], pp. 139-40. Coefficients require 1 for application of this method are tabulated in Abramowitz and Stegan [1972].

The Pascal source code given by Press et al was used as a prototype for the routine developed here, however modifications of this routine were required. The first of these entailed development of a seperate routine to initialize arrays with the abscissca and weight values required by the Gauss-Legendre method. These arrays were defined as global variables in the program. This stands in sharp contrast to the algorithm given by Press et al [1989], who re-initialize these values each time the routine is called. By carrying these values as global variables a considerable gain in run-time performance has been attained at the cost of a small increase in memory requirements of the program. It was also necessary to cast this procedure in a nested loop form, one loop for each integral, each loop step entailing four evaluations of the integration function.

Calculation of the free energy per unit area is achieved in a single program loop in the main program body. Each pass through the loop corresponds to calculation of a single frequency term in the frequency sum appearing in eqn. (4.2.15). The dielectric permittivity corresponding to each frequency is calculated each time the loop is executed, as is the corresponding contribution to the free energy. This latter is accomplished by calls to the aforementioned

integration routines. The absolute value of the ratio of the last term calculated to the cumulative sum of terms is calculated after each pass through the loop and tested against a predefined value to determine when enough terms have been summed.

An output procedure was incorporated into the program to write an output file summarizing the results for a given adsorbed layer thickness and range of capillary radii. A utility program was then written to read this file, calculate comparision values based on the work of Philip [1977b], and generate a report file which lists these values. The data tables provided in appendix D have been taken from these report files.

4.4 Results and Conclusions

4.4.1 Introductory Comments

Gold and silver were chosen as substrate materials for this study because extensive tabulations of optical data are available for both materials and in consideration of the fact that both metals form relatively clean surfaces. Data files containing the optical properties of gold and silver were created on the basis of the data tabulated in Palik [1985]. This source, which is based on an extensive review of experimental results prior to 1985, lists the optical properties for a number of metals over a large range of incident photon energies. The data file for gold consists of 271 points spanning and energy range of 0.125 - 87.50 eV. Similiarly, the data file for silver contains 238 points covering an energy range from 0.125 - 10,000 eV. The optical data files for both gold and silver are listed in appendix E of this work.

Pentane and octane were considered in this study because

these liquids have been studied extensively.
Ninham-Parsegian representations developed by Hough and
White [1980] were used to model the dielectric
permittivities of the cor᠎densed liquid phase. These
representations, which fi᠎ experimental results quite well,
are based on one IR absorption peak and one UV absorption
peak. Hough and White have pointed out that microwave
absorption peaks only make a significant contribution when
calculating dispersion forces at very low temperatures. The
specific number of peaks required to adequately represent a
given liquid varies. Water, for example, is typically
represented with between five and eight absorption peaks in
the Ninham-Parsegian representation.

The percent deviation of the van der Waals free energy
of interaction in a capillary tube from that in a flat plate
geometry characterizes the effect of capillary curvature on
the van der Waals forces. This quantity is defined as

$$\delta_{surf} = 100 \frac{F_{tube} - F_{plate}}{F_{plate}},$$ (4.4.1)

where F_{plate} denotes the free energy calculated for a flat
plate geometry with a given condensed liquid phase thickness
and F_{tube} is the corresponding value for a capillary tube of
a given radius. The subscript surf is used to indicate that
this result was obtained using the surface mode analysis
method. For comparison purposes, it is possible to obtain a
corresponding expression for the Hamaker theory from the
work of Philip [1977b]. To first order in l/r_c this is given
by

$$\delta_{Ham} = 150 \frac{l}{r_c},$$ (4.4.2)

The ratio of the quantities defined by eqns. (4.4.1) and
(4.4.2) gives an important measure of the extent to which
the predictions of the Hamaker theory and the surface mode

analysis methods coincide. These quantities are all tabulated in each table of results in appendix F.

Results were obtained over a range of capillary radii from 5 - 60 microns in increments of five microns, and a range of condensed liquid phase thicknesses from 10 to 90 nm in increments of 20 nm. All results given were calculated for a temperature of 300 K. In obtaining the results, the limits of integration of eqn. (4.2.23) were varied for each run until the calculated free energies showed no change in the first six decimal places. All computer calculations were run on a Compudyne 25/386c 25 MHz 80386 based personal computer with an IIT math co-processor. A typical program run took approximately 15 minutes.

For each run a data file was generated listing the condensed liquid phase, the substrate, the thickness of the condensed liquid phase, and the free energy per unit area for each capillary radius value. The results which have been tabulated in appendix F were generated from these files by use of a conversion program written for that purpose. The conversion program generates an output file which tabulates both the surface mode analysis and Hamaker theory predictions.

4.4.2 Remarks on the Results

Calculations carried out as described in section 4.1 are summarized in figs. 4.1-4.4 and in appendix F. These results merit several remarks. First of all, the magnitude of the van der Waals free energy per unit area calculated for pentane and octane on gold compares favorably with those calculated from the work of Wayner [1978] for thin condensed liquid phases. A reasonable comparison can be made for $l = 100 A$, where Wayner's results predict that F_{plate} =

1.32659e-5 J/m^2 for pentane on gold and F_{plate} = 1.39261e-5 J/m^2 for octane on gold. These values are higher than those given in Tables D.1 and D.11 by 1.63% and 1.56%, respectively. This deviation may be accounted for by three facts: (1) The work of Wayner does not include temperature dependence, and in fact is based on a value at absolute zero, (2) Wayner's results also neglect the effect of retardation on the van der Waals interaction, and (3) the present work is based on better optical data than was available in 1978. In addition, some of the deviation may be due to the particular numerical methods used to calculate the results.

Inclusion of retardation effects constitutes the single most important difference between the present work and efforts based on the Hamaker theory. In fig. 4.1 a comparison has been made of the magnitude of the van der Waals forces as predicted by the Hamaker theory using the results of Wayner [1978] and the present work. The Hamaker theory over-estimates the magnitude of these forces by as much as 50% over the range of thicknesses of interest.

In the present work, temperature effects were found to be rather small. Fig. 4.2 is a plot of the percent deviation of the van der Waals free energy per unit area at a given temperature from that calculated at 300 K. These results, which are typical, have been calculated for a 10 nm thick condensed liquid phase of pentane on a gold substrate. Over a span of 100 K, the van der Waals free energy of interaction per unit area was found to vary by 0.2965% from the computed value at 300 K. Similiar results are obtained for other condensed liquid phases and substrate materials. The van der Waals interaction decreases with temperature because thermal excitation tends to disalign dipoles within the system. Although the van der Waals free energy per unit area varies relatively little with temperature, it is

important to include temperature dependence in calculations of this interaction. Omission of temperature dependence can result in over-predicting the magnitude of the van der Waals interaction by as much as 1.0%, which is significant. This topic has been pursued in some detail by Parsegian and Ninham [1970].

The effect of capillary curvature on the van der Waals interaction in capillary tubes can clearly be seen in fig. 4.3. This plot shows the percent deviation of the van der Waals interaction in a capillary tube from that in a flat plate geometry. This deviation asymptotes to zero for large capillary radii irrespective of condensed liquid phase layer thickness, as one would expect. In the limit of small capillary radii, this deviation increases sharply, especially for thick condensed liquid phases. On this basis, one would expect enhanced capillary pumping for small capillary radii and thick condensed liquid phases. The significance of this observation is that surface roughness effects and effects due to surface impurities should be negligible in the thick film region. Consequently, the effect of capillary curvature constitutes the dominant correction to the Lifshitz theory (Appendix C) in this regime.

It is notable that the trends predicted by Philip [1977b] and the present work for the effect of capillary curvature also compare quite favorably. This can be seen in fig. 4.4, which is a plot of $\delta_{surf}/\delta_{Ham}$ vs r_c. The reader is cautioned that the actual magnitudes predicted by the surface mode analysis method and by the work of Philip [1977b] will be substantially different because Philip has not accounted for retardation. The ratio of deviations tends to be within a few percent of 1.0. For small radii and thin films, application of the Hamaker theory is reasonable. The results presented indicate that the surface mode analysis

method predicts a greater curvature effect than that
predicted by the Hamaker theory. The strong dependence of
$\delta_{surf}/\delta_{Ham}$ on the thickness of the condensed liquid phase
suggests that deviations of the surface mode analysis
results from those predicted by Philip [1977b] are due to
retardation effects.

4.4.3 Conclusions

The present work has refined previous efforts to
evaluate van der Waals forces in capillary tubes. This
effort constitutes an essential step in evaluating the role
of van der Waals forces in capillary pumping. Preliminary
results suggest that these forces do in fact play a
significant role and thus that heat pipe wick performance
might well be enhanced by appropriate materials selection
and processing. It is clear, however, that these effects
have been overestimated in past works based on the Hamaker
theory.

The value of the present work lies in the fact that it
accounts for the temperature dependence and retardation of
van der Waals forces in capillary tubes. In addition, the
model presented here is based on a more sophisticated
desciption of material properties than previous efforts
based on the Hamaker theory. The fact that capillary
curvature effects are enhanced for thick condensed liquid
phases indicates that these effects will be probably play a
significant role in this region, surpassing effects due to
surface roughness and surface impurities, such as oxide
layers. Unfortunately, a specific and proper comparision
with surface roughness and curvature effects is not possible
at this time, because their specific effect on interactions

with an adsorbed layer systems have not been treated. Estimation of these effects must thus be made on the basis of very general studies.

4.5 Nomenclature

Roman

$\vec{A}$ - Vector potential

$\vec{a}$ - Fourier coefficient in superpotential

$\vec{b}$ - Fourier coefficient in superpotential

c - Speed of light

$\vec{c}$ - Fourier coefficient in superpotential

D - Dispersion equation

$\vec{d}$ - Fourier coefficient in superpotential

$\vec{e}$ - Fourier coefficient in superpotential

F - Van der Waals free energy per unit area

$\vec{f}$ - Fourier coefficient in superpotential

K_B - Boltzmann constant

k - Wave number for surface modes

l - Thickness of condensed liquid phase

m - Index for theta modes

n - Index for electron frequencies

p - Nondimensional wave number for surface modes

r_c - Radius of capillary tube

r_i - Radial location of vapor-liquid interface

T - Temperature (absolute)

$\vec{Z}$ - Superpotential

Greek

$\bar{\Delta}$ - Coefficients in normal dispersion equation

Δ - Coefficients in tangential dispersion equation

δ - Percent deviation from the flat plate limit

ϵ - Dielectric perm.ttivity

μ - Magnetic permeability

ω - Frequency of surface modes

ϕ - Scalar potential

ρ - $\sqrt{k^2 - \mu\epsilon\omega^2\left(1 + \frac{i o}{\epsilon\omega}\right)}$

σ - Electrical conductivity

ξ_n - $2\pi n K_B T / \hbar$

ξ_s - $2l\sqrt{\epsilon_2}/c$

Subscipts

ω - Denotes Fourier component

m - Refers to m-th theta mode

r - Denotes radial vector component

θ - Denotes theta vector component

z - Denotes longitudinal vector component

1 - Pertains to capillary tube medium

2 - Pertains to vapor phase

3 - Pertains to condensed liquid phase

5. RECOMMENDATIONS

There are many aspects of this research that require further investigation. The first is testing the model over a wide range of parameters for different nonpolar working fluids (i.e. cryogenic fluids). These tests will be conducted shortly after this paper is complete. Secondly, all of the results in this work require experimental verification. Unfortunately, performing an experiment on a capillary tube with a radius less than 100 microns is a formidable task. It may be possible to set up an experiment with a bundle of capillary tubes to allow the measurement of an effective mass transfer rate. This would require more extensive modeling to account for the variation in tube wall temperature between the outer wall and the bundle centerline. Similar tests on tubes, or vertical flat plates, with vapor deposited surface layers of varying thickness and material type should clarify the effect surface coatings have on meniscus pumping capability.

Much theoretical work remains to be done with regard to calculation of the van der Waals forces in capillary tubes. Although the present work has remedied many of the deficiencies of treatments based on the Hamaker theory, a clear need exists for further refinement of the present model to account for the properties of real surfaces. Factors which might be considered in extending this work include the presence of chemical impurities on the surface of the capillary tube, surface roughness effects, and electric double-layer forces. The presence of surface charge concentrations can lead to coupling of the van der Waals forces and the electric double-layer forces. This point should be

throughly explored. It should be emphasized that all of
these corrections are only relevant for condensed
liquid phases of thicknesses less than 20 nm or so.

6. REFERENCES

Abramowitz, M., and Stegan, I.A.,[1972], <u>Handbook of</u> <u>Mathematical Functions</u>, Dover Publications, Inc., New York.

Abrikosov, A.A., Gor'kov, L.P., and Dzyaloshinskii, I.E.,[1963], <u>Methods of Quantum Field Theory in Statistical</u> <u>Physics</u>, Prentice-Hall, Englewood Cliffs, New York.

Adamson, A.W., and Zebib, A., [1980], "Transition Region Between an Infinite Plane Meniscus and an Absorbed Film," Journal of Phys. Chem., V84 pp. 2619-2623.

Adamson, A.W., [1982], <u>Physical Chemistry of Surfaces</u>, 4th Edition, John Wiley and Sons, New York.

Ambrose, J.H., Chow, L.C., and Beam, J.E., [1987],"Transient Heat Pipe Response and Rewetting Behavior," Journal of Thermophysics and Heat Transfer, V1, No.3, pp. 222-227.

Barash, Yu.S., and Ginzburg, V.L.,[1989], "Electromagnetic Fluctuations and Molecular Forces in Condensed Matter," in <u>The Dielectric Function of Condensed Systems</u>, Edited by Keldysh,L.V., Kirzhnitz, D.A., and Maradudin, A.A.,North-Holland Pub. Co., New York, pp. 389-457.

Barton, G.,[1979], "Some Surface Effects in the Hydrodynamic Model of Metals," Rep. Prog. Phys., Vol. 42, pp. 963-1016.

Beam, J.E., [1985], "Unsteady Heat Transfer in Heat Pipes," Ph.D. Dissertation, School of Engineering, University of Dayton.

Beam, J.E., [1985], "Transient Heat Pipe Analysis," AIAA Paper 85-0936.

Burden, R.L., and Faires, J.D.,[1985], <u>Numerical Analysis</u>, 3rd ed., Prindle, Weber, and Schmidt, Boston.

Casimir,H.B.G and Polder, D.,[1948], "The Influence of Retardation on the London-van der Waals Forces," Phys. Rev., Vol. 73 No. 4, pp. 360-372.

Davies, B.,[1971], "Extension of van Kampen et al.'s Method of Calculating van der Waals Forces to Dissipative Media," Phys. Lett., Vol. 37A, No. 5, pp. 391-392.

Deryagin, B.V., Newpin, S.V., and Churayev, N.V., [1965], "Effect of Film Transfer Upon Evaporation of Liquids from Capillaries," Bull. R.I.L.E.M., V29, pp. 93-98.

Dzyaloshinkskii,I.E., Lifshitz, E.M., and Pitaevskii, L.P.,[1960], "Van der Waals Forces in Liquid Films," Sov. Phys. JETP., Vol. 37, No. 1, pp. 161-170.

Dzyaloshinskii, I.E., Lifshitz, E.M., and Pitaevskii, L.P.,[1961], "The General Theory of van der Waals Forces," Adv. Phys., Vol. 10, pp. 165-209.

Gingell, D., and Parsegian, V.A.,[1972], "Computation of van der Waals Interactions in Aqueous Systems Using Reflectivity Data," J. Theor. Biol., Vol. 36, pp. 41-52.

Gerlach, E.,[1971], "Equivalence of van der Waals Forces Between Solids and the Surface-Plasmon Interaction," Phys. Rev. B., Vol. 4, No. 2, pp. 393-396.

Haines, W.B., [1930], "Studies in the Physical Properties of Soils, Part V," Journal of Agric. Sci. Camb..., V20, pg. 97.

Hamaker, H.C.,[1937], "The London-van der Waals Attraction Between Spherical Particles," Physica, Vol. 4, pp. 1058-1072.

Hazlett, R.D.,[1990], "Fractal Applications: Wettibility and Contact Angle," J. Coll. Int. Sci., Vol. 137, No. 2, pp. 527-533.

Herman, M.C., and Papadopoulis, K.D.,[1990],"Effects of Asperities on the van der Waals and Electric Double-layer Interactions of Two Parallel Flat Plates," J. Coll. Int. Sci., Vol. 136, No. 2, pp. 385-392.

Hill, T.L., [1949], "Statistical Mechanics of Adsorption. V. Thermodynamics and Heat of Adsorption," Journal of Chem. Phys., V17, No. 6, pp. 520-535.

Hill, T.L., [1949], "Thermodynamic Transition from Adsorption to Solution," Journal of Chem. Phys., V84, No. 6, pg. 507.

Hill, T.L., [1952], "Theory of Physical Adsorption", in Advances in Catalysis, IV, Academic Press, pp. 211-258.

Horn, R.G., and Israelachvili, J.N.,[1981],"Direct Measurement of Structural Forces Between Two Surfaces in a Nonpolar Liquid," J. Chem. Phys., Vol. 75, No. 3, pp. 1400-1411.

Hough, D.B., and White, L.R.,[1980], "The Calculation of Hamaker Constants from Lifshitz Theory with Applications to Wetting Phenomena," Adv. Coll. Int. Sci., Vol. 14, pp. 3-41.

Israelachvili, J.N., [1985], <u>Intermolecular and Surface Forces with Application to Colloidal and Biological Systems</u>, Academic Press, New York.

Israelachvili, J.N, and Tabor, D.,[1972], "The Measurement of the van der Waals Dispersion Forces in the Range 1.5 to 130 nm," Proc. Roy. Soc. A, Vol. 331, pp. 19-38.

Krupp, H., Schnabel, W. and Walter, G.,[1972], "Computation of the Lifshitz-van der Waals Constant on the Basis of Optical Data," J. Coll. Int. Sci., Vol. 39, No. 2, pp. 421-423.

Ivanov, I.B., [1988], <u>Thin Liquid Films</u>, Marcel Dekker, Inc., New York.

Landau, L.D., and Lifshitz, E.M.,[1980], <u>Statistical Physics</u>, Part 1, 3rd ed. Pergamon Press , Oxford, England.

Langbein, D.,[1974], <u>Theory of van der Waals Attraction</u>, Springer-Verlag, New York.

Lifshitz, E.M.,[1955], "Th Theory of Molecular Attractive Forces Between Solids," Sov. Phys. JETP, Vol. 2, No. 1, pp. 73-83.

Mahanty, J. and Ninham, B.W.,[1976], <u>Dispersion Forces</u>, Academic Press, New York.

Maradudin, A.A., and Mazur, P.,[1980], "Effects of Surface Roughness on the van der Waals Force Between Macroscopic Bodies," Phys. Rev. B., Vol. 22, No. 4, pp. 1677-1686.

Marvin, A.M., and Toigo, F.,[1982],"Van der Waals Interaction Between a Point Particle and a Metallic Surface. I. Theory.," Phys. Rev. A., Vol. 25, No. 2, pp. 782-802.

Meyer, R.E., [1984], "Note on Evaporation in Capillaries," IMA Journal of Applied Mathematics, V32, pp. 236-252.

Mirzamoghadam, A., and Catton, I., [1988], "A Physical Model of the Evaporating Meniscus," Journal of Heat Transfer, V110, No. 1, pp. 201-207.

Moon, P. and Spencer, D.E.,[1988], <u>Field Theory Handbook</u>, 3rd ed., Springer-Verlag, New York.

Ninham, B.W., Parsegian, V.A. and Weiss, G.H,[1970], "On the Macroscopic Theory of Temperature-Dependent van der Waals Forces," J. Stat. Phys., Vol. 2, No. 4, pp. 323-328.

Ninham, B.W. and Parsegian, V.A.,[1970a],"Van der Waals Forces Across Triple-Layer Films," J. Chem. Phys., Vol. 52, No. 9, pp. 4578-4587.

Ninham, B.W. and Parsegian, V.A.,[1970b], "Van der Waals Forces : Special Charecteristics in Lipid-Water Systems and a General Method of Calculation Based on the Lifshitz Theory," Biophys. J., Vol. 10, pp. 646-663.

Nir, S., Rein, R., and Weiss, L.,[1972], "On the Applicability of Certain Approximations of the Lifshitz Theory to Thin Films," J. Theor. Bio., Vol. 34, pp. 135-153.

Nir, S. and Vassilieff, C.S.,[1988],"Van der Waals Interactions in Thin Films," in Thin Liquid Films, Edited by Ivanov, I.B., Marcel Dekker, Inc., New York, New York, pp. 207-274.

Palik, E.D., ed.,[1985], Handbook of Optical Constants of Solids, Academic Press, Inc. New York.

Panofsky, W., and Phillips, M.,[1962], Classical Electricity and Magnetism, 2nd ed., Addison-Wesley Publishing Co., New York.

Parsegian, V.A., and Ninham, B.W.,[1970], "Temperature-Dependent van der Waals Forces," Biophys. J., Vol. 10, pp. 664-674.

Pashley, R.M.,[1977], "The van der Waals Interaction for Liquid Water: A Comparison of the Oscillator Model Approximation and Use of the Kramers-Kronig Equation with Full Spectral Data," J. Coll. Int. Sci., Vol. 62, No. 2, pp. 344-347.

Philip, J.R., [1977a], "Unitary Approach to Capillary Condensation and Adsorption," Journal of Chemical Physics, V66, No. 11, pp. 5069-5075.

Philip, J.R.,[1977b], "Adsorption and Geometry: the Boundary Layer Approximation," J. Chem. Phys., Vol. 67, No. 4, pp. 1732-1741.

Philip, J.R.,[1978], "Adsorption and Capillary Condensation on Rough Surfaces," J. Phys. Chem., Vol. 82, No. 12, pp. 1379-1385.

Preiss, G., and Wayner, P.C. Jr., [1976], "Evaporation from a Capillary Tube," Journal of Heat Transfer, V98, No. 2, pp. 178-181.

Press, W.H., et al,[1989], <u>Numerical Recipes in Pascal, The Art of Scientific Computing</u>, Cambridge University Press, Cambridge, England.

Rahman, T.S., and Maradudin, A.A.,[1980], "Effect of Surface Roughness on the Image Potential," Phys. Rev. B., Vol. 21, No. 2, pp. 504-521.

Richmond, P., and Ninham, B.W.,[1971a], "A Note on the Extension of the Lifshitz Theory of van der Waals Forces to Magnetic Media," J. Phys. C., Vol. 4, pp. 1988-1993.

Richmond, P., and Ninham, B.W.,[1971b] "Calculations of van der Waals Forces Across Films of Liquid Helium Using Lifshitz Theory," J. Low Temp. Phys., Vol. 5 No. 2, pp. 177-189.

Richmond, P., Ninham, B.W., and Ottewill,R.H.,[1973], "A Theoretical Study of Hydrocarbon Adsorption on Water Surfaces Using Lifshitz Theory," J. Coll. Int. Sci., Vol. 4`, No. 1, pp. 69-80.

Richmond, P.,[1975], "The Theory and Calculation of van der Waals Forces," in <u>Colloid Science: A Specialist Periodical Report</u>, The Chemical Society, London.

Sabisky, E.S., and Anderson, C.H.,[1973],"Verification of the Lifshitz Theory of the van der Waals Potential Using Liquid-Helium Films," Phys. Rev. A, Vol. , pp. 790-

Schrage, R.W., [1953], <u>A Theoretical Study of Interphase Mass Transfer</u>, Columbia University Press, New York.

Swanson, L.W., and Herdt, G.C.,[1990], "Model of the Evaporating Meniscus in a Capillary Tube," Accepted by the 3rd ASME-JSME Thermal Engineering Joint Conference.

van Bree, J.L.M.J, et al,[1974], "The Influence of Surface Irregularities upon the van der Waals Forces Between Macroscopic Bodies," Physica, Vol. 78, pp. 187-190.

van Kampen, N.G., Nijboer,B.R.A., and Schram, K.,[1968], "On the Macroscopic Theory of van der Waals Forces," Phys. Lett., Vol. 26A, No. 7, pp. 307-308.

Vargaftik, N.B., [1975], <u>Tables on the Thermophysical Properties of Liquids and Gases</u>, 2nd Edition, John Wiley and Sons, Inc., New York.

Verwey, E.J., and Overbeek, J. Th. G.,[1948], <u>Theory of the Stability of Lyophobic Colloids</u>, Elsevier, New York.

Wayner, P.C. Jr., Kao, Y.K., and LaCroix, L.V., [1976], "The Interline Heat-Transfer Coefficient of an Evaporating Wetting Film," International Journal of Heat and Mass Transfer, V19, pp. 487-492.

Wayner, P.C. Jr., [1978], "The Effect of the London-Vander Waals Dispersion Force on Interline Heat Transfer," Journal of Heat Transfer, V100, No. 1, pp. 155-159.

Wayner, P.C. Jr., [1979], "Effect of Thin Film Heat Transfer on Meniscus Profile and Capillary Pressure," AIAA Journal, V17, No. 7, pp. 772-776.

Wayner, P.C. Jr., [1986], "The Use of Interfacial Phenomena in Change-Of-Phase Heat Transfer," Heat Transfer in Thermal Systems Seminar-Phase II, National Cheng Kung University, Tainan, China.

Weast, R.C., ed.,[1981], CRC Handbook of Chemistry and Physics, 60th ed., CRC Press, Inc., Boca Raton, Florida.

7. Appendix A

The dimensional form of the mathematical functions which were generated during the formulation of the problem in Chapter 3 are:

$$F_1 = \frac{\frac{8\mu_l}{\pi\rho_l}}{-3r_i^4 + 4r_c^2 r_i^2 - r_c^4 + r_i^4 \ln\left(\frac{r_i}{r_c}\right)}, \qquad (A1)$$

$$F_2 = \frac{4\left(r_i^3 - r_c^3 r_i - 2r_i^3 \ln\left(\frac{r_i}{r_c}\right)\right)}{-3r_i^4 + 4r_c^2 r_i^2 - r_c^4 + r_i^4 \ln\left(\frac{r_i}{r_c}\right)}, \qquad (A2)$$

$$F_3 = -\frac{1}{4}(r_c^2 - r_i^2)^2 - \frac{-\frac{2}{5}r_c^5 + \frac{1}{20}r_i^5 + \frac{1}{4}r_c^4 r_i - \frac{1}{6}r_c^2 r_i^5}{r_c - r_i}, \qquad (A3)$$

$$F_4 = T_i(-r_i^3 + r_c^2 r_i) + \frac{1}{4}\frac{T_w - T_i}{r_c - r_i}(r_c^2 - r_i^2)^2$$

$$- \frac{T_w - T_i}{(r_c - r_i)^2}\left(-\frac{2}{15}r_c^5 + \frac{1}{20}r_i^5 + \frac{1}{4}r_c^4 r_i - \frac{1}{6}r_c^2 r_i^3\right), \qquad (A4)$$

$$F_5 = -\frac{1}{2}r_i^2 \ln\left(\frac{r_i}{r_c}\right) - \frac{1}{4}(r_c^2 - r_i^2)$$

$$- \frac{\frac{1}{6}r_i^3 \ln\left(\frac{r_i}{r_c}\right) - \frac{1}{9}r_c^3 + \frac{1}{4}r_c^2 r_i - \frac{5}{36}r_i^3}{r_c - r_i}, \qquad (A5)$$

$$F_6 = -T_i r_i \ln\left(\frac{r_i}{r_c}\right) + \frac{T_w - T_i}{r_c - r_i}\left\{\frac{1}{2}r_i^2 \ln\left(\frac{r_i}{r_c}\right) + \frac{1}{4}(r_c^2 - r_i^2)\right\}$$

$$- \frac{T_w - T_i}{(r_c - r_i)^2}\{\frac{1}{2}r_i^3 \ln\left(\frac{r_i}{r_c}\right) + \frac{1}{4}r_c^2 r_i - \frac{1}{4}r_i^3$$

$$- \frac{1}{3}r_i^3 \ln\left(\frac{r_i}{r_c}\right) - \frac{1}{9}(r_c^3 - r_i^3)\}, \qquad (A6)$$

The nondimensional mathematical functions) found in the dimensionless working equation set (3.19-3.23) are:

$$G_1 = \frac{1}{-3\bar{r}_i^4 + 4\bar{r}_i^2 - 1 + 4\bar{r}_i^4 \ln \bar{r}_i},$$

(A7)

$$G_2 = \left(\bar{r}_i^3 - \bar{r}_i - \bar{r}_i^3 \ln \bar{r}_i\right) G_1,$$

(A8)

$$G_3 = -\frac{1}{4}\left(1 - \bar{r}_i^2\right)^2 - \frac{-\frac{2}{5} + \frac{1}{20}\bar{r}_i^5 + \frac{1}{4}\bar{r}_i - \frac{1}{6}\bar{r}_i^3}{1 - \bar{r}_i},$$

(A9)

$$G_4 = \frac{T_w}{T_w - T_v}\left(-\bar{r}_i^3 + \bar{r}_i\right) - \left(-\bar{r}_i^3 + \bar{r}_i\right)\bar{T}_i - \frac{1}{4}\frac{\left(1 - \bar{r}_i^2\right)^2}{1 - \bar{r}_i}\bar{T}_i$$
$$+ \frac{-\frac{2}{5} + \frac{1}{20}\bar{r}_i^5 + \frac{1}{4}\bar{r}_i - \frac{1}{6}\bar{r}_i^3}{\left(1 - \bar{r}_i\right)^2}\bar{T}_i,$$

(A10)

$$G_5 = -\frac{1}{2}\bar{r}_i^2 \ln \bar{r}_i - \frac{1}{4}\left(1 - \bar{r}_i^2\right) - \frac{\frac{1}{6}\bar{r}_i^3 \ln \bar{r}_i - \frac{1}{9} + \frac{1}{4}\bar{r}_i - \frac{5}{36}\bar{r}_i^3}{1 - \bar{r}_i},$$

(A11)

$$G_6 = -\frac{T_w}{T_w - T_v}\bar{r}_i \ln \bar{r}_i + \bar{r}_i \bar{T}_i \ln \bar{r}_i + \frac{\frac{1}{2}\bar{r}_i^2 \ln \bar{r}_i + \frac{1}{4}\left(1 - \bar{r}_i^2\right)}{1 - \bar{r}_i}\bar{T}_i$$
$$- \frac{\frac{1}{6}\bar{r}_i^3 \ln \bar{r}_i - \frac{1}{9} + \frac{1}{4}\bar{r}_i - \frac{5}{36}\bar{r}_i^3}{\left(1 - \bar{r}_i\right)^2}\bar{T}_i,$$

(A12)

$$G_7 = \left(\frac{T_w}{T_w - T_v} - \bar{T}_i\right)^{1/2},$$

(A13)

8. Appendix B

The program used to simulate equations (3.19)-(3.23) in Chapter 3 describing the evaporating meniscus in a capillary tube is listed below.

```
C
C ---------- THE EVAPORATING MENISCUS IN A CAPILLARY TUBE ----------
C
C
C                      by
C
C
C              Larry W. Swanson
C
C
C
C  Background: This program simulates the evaporating meniscus in a capillary
C  tube. A system of five nonlinear ordinary differential equations are solved
C  by initially quasi-linearizing the first order terms. The resulting
C  equation set takes the form of a fifth order linear matrix equation. An
C  implicit scheme is used to solve the matrix equation which updates the
C  coefficient matrix during each iteration. The solution vector contains the
C  five unknown spatial derivatives of the dependent variables. The dependent
C  variables are evaluated during each iteration using a first order backward
C  finite difference approximation for their derivatives. Once the dependent
C  variables and their derivatives converge to within a given relative error
C  tolerance the program moves forward to the next spatial location.
C
C
      PROGRAM CAPMAIN
C
C ---------- Initialize derivatives and dependent variables ----------
C
C  Variable list (dimensionless):
C
C  X - dependent variable vector for the N + 1th iteration.
C
C      X(1) = ri (dimensionless meniscus radial location)
C      X(2) = Ti (dimensionless interfacial temperature)
C      X(3) = mdot (dimensionless mass transfer rate)
C      X(4) = Pli (dimensionless liquid interfacial pressure)
C      X(5) = dri/dx = Di (first derivative of the meniscus radial location)
C  or  X(5) = w ri (transformed coordinate, see w below)
C
C  w - 1/dsqrt(1 + Di**2)
C
C  XIT - dependent variable vector for the Nth iteration.
C
C  Y - the 5x7 equation set matrix. The first five columns represent the
C      coefficients while the sixth and seventh columns represent the
C      dX(i)/dx and the constant vector.
C
C  YIT - derivative of the variable vector (Y(N,6)) for the Nth iteration.
C
C  XLOC - current x location.
C
C  STEP - step size.
C
C  NSTEPS - number of steps.
C
C  ERROR - convergence relative error tolerance.
C
C
C ---------------- Dimension statements ----------------
C
      DIMENSION Y(5,7),YIT(5),Y0(5),X(5),XIT(5),X0(5),PI(6)
      IMPLICIT REAL*8 (A-H,O-Z)
      REAL*8 MUL,MW,KL
      N = 5
      NN = 7
      XLOC = 0.D0
      STEP = 0.0001D0
      NSTEPS = 10000
      ERROR = 1.D-08
      TPOINT = -1.D-03
      NFLAG = 0
      OPEN(5,FILE = 'CAPDATA.DAT',STATUS = 'OLD')
```

48-76

```
C
C ——————— Console input parameters ———————
C
C Input Parameters (SI units):
C
C ABAR - dispersion coefficient.
C
C DELT - superheat.
C
C RC - capillary tube radius.C
C
C
    WRITE(5,1)
  1 FORMAT('Test Data for the Captube Program')
    PRINT2
  2 FORMAT(6X,'Abar (1PD9.2) = ')
    READ(*,3)ABAR
  3 FORMAT(1PD9.2)
    PRINT4
  4 FORMAT(6X,'Superheat (1PD9.2) = ')
    READ(*,5)DELT
  5 FORMAT(1PD9.2)
    PRINT6
  6 FORMAT(6X,'Capillary radius (1PD9.2) = ')
    READ(*,7)RC
  7 FORMAT(1PD9.2)
C
C ——————— Define fluid and material properties (hexane at 290 K) ———————
C
C Material Properties (SI units):
C
C TV - saturated vapor temperature.
C
C PV - saturated vapor pressure.
C
C TW - wall temperature.
C
C RHOL - liquid density.
C
C CP - liquid specific heat.
C
C KL - liquid thermal conductivity
C
C ALPHA - thermal diffusivity.
C
C HFG - latent heat of vaporization.
C
C SIGMAW - liquid surface tension.
C
C DSIGDT - derivative of the surface tension with respect to temperature.
C
C MUL - liquid viscosity.
C
C C - vapor-liquid accommodation coefficient.
C
C MW - liquid molecular weight.
C
C R - ideal gas constant divided by the molecular weight.
C
C PIE - 3.14159.......
C
C
    C = 1.0D0
    MW = 86.172D0
    R = 8.3144D0/MW
    PIE = 3.14259D0
C
C High temperature data for hexane
C
C    TV = 453.D0
C    PV = 1.307D06
C    TW = TV+DELT
C    RHOL = 475.1D0
C    CP = 2270.D0
C    KL = 0.0300D0
C    ALPHA = KL/RHOL/CP
C    HFG = 214.3D03
C    SIGMAW = 3.16D-03
C    DSIGDT = -7.450D-05
C    MUL = 811.D-07
C
```

```fortran
C Room temperature data for hexane
C
      TV = 290.D0
      PV = 1.6163D04
      TW = TV+DELT
      RHOL = 659.4D0
      CP = 2239.D0
      KL = 0.126D0
      ALPHA = KL/RHOL/CP
      HFG = 368.3D03
      SIGMAW = 18.46D-03
      DSIGDT = -1.055D-04
      MUL = 3070.D-07
C
C Low temperature data for hexane
C
C     TV = 263.D0
C     PV = 3.437D03
C     TW = TV+DELT
C     RHOL = 686.2D0
C     CP = 2134.D0
C     KL = 0.138D0
C     ALPHA = KL/RHOL/CP
C     HFG = 384.3D03
C     SIGMAW = 21.61D-03
C     DSIGDT = -1.050D-04
C     MUL = 4260.D-07
C
C ------ Determine equation parameters or dimensionless numbers {PIi} ------
C
      PI(1) = RC*PV/4.D0/SIGMAW
      PI(2) = 2.D0*MUL*KL*DELT/PIE/RC/RHOL/SIGMAW/HFG
      PI(3) = ALPHA*MUL/RC/SIGMAW
      PI(4) = (2.D0-C)/C*R**(5)*KL*DELT**(1.5)
      PI(4) = PI(4)/2.D0**(.5)/PIE**(.5)/SIGMAW/HFG
      PI(5) = -ABAR/SIGMAW/RC/RC
      PI(6) = -DSIGDT*DELT/SIGMAW
      SIGBAR = 1.D0
C
C ------ Determine the initial values of the dependent variables {X0(i)} ------
C
      X0(1) = -ABAR*TV/RHOL/HFG/DELT/RC/RC/RC
      X0(1) = 1.D0-X0(1)**(1.D0/3.D0)
      X0(2) = 2.D-11
      X0(3) = -2.D-11
      X0(4) = 1.D0-SIGMAW/PV/RC/X0(1)-RHOL*HFG*DELT/PV/TV
      X0(5) = -2.D-11
C
C ------------- Initialize the dependent variables {X(i)} -------------
C
      X(1) = X0(1)
      X(2) = X0(2)
      X(3) = X0(3)
      X(4) = X0(4)
      X(5) = X0(5)
      DO 8 L=1,5
      XIT(L) = X(L)
    8 CONTINUE
C
C ------ Initialize the derivatives of the dependent variables {Y(i)} ------
C
      Y(1,6) = -5.D-08
      Y(2,6) = 5.D-08
      Y(3,6) = -5.D-08
      Y(4,6) = 5.D-08
      Y(5,6) = -5.D-08
      DO 9 L=1,5
      YIT(L) = Y(L,6)
    9 CONTINUE
C
C ------------- Write parameters and initial values into 'capdata' file -------------
C
      WRITE(5,10)NSTEPS
   10 FORMAT(2X,'Number of steps = ',I5)
      WRITE(5,11)ABAR
   11 FORMAT(2X,'Abar = ',1PD9.2)
      WRITE(5,12)DELT
   12 FORMAT(2X,'Superheat = ',1PD9.2)
      WRITE(5,13)RC
   13 FORMAT(2X,'Capillary radius = ',1PD9.2)
      WRITE(5,14)TPOINT
```

```
   14 FORMAT(2X,'Transition point = ',1PD9.2)
      DO 16 K=1,6
      WRITE(5,15) K,PI(K)
   15 FORMAT(2X,'PI',I1,2X,' = ',1X,1PE13.6)
   16 CONTINUE
      WRITE(5,17)
   17 FORMAT('Initial values of the dependent variables')
      DO 19 K=1,5
      WRITE(5,18) K,X0(K)
   18 FORMAT(2X,'X0',I1,2X,' = ',1X,1PE13.6)
   19 CONTINUE
      WRITE(5,20)
   20 FORMAT('Initial guesses for the dependent variables')
      DO 22 K=1,5
      WRITE(5,21) K,X(K)
   21 FORMAT(2X,'X',I1,2X,' = ',1X,1PE13.6)
   22 CONTINUE
      WRITE(5,23)
   23 FORMAT('Initial guesses for the derivatives of the dependent varia
     *bles')
      DO 25 K=1,5
      WRITE(5,24) K,Y(K,6)
   24 FORMAT(2X,'Y',I1,2X,' = ',1X,1PE13.6)
   25 CONTINUE
C
C ------------------ Begin Program Calculations ----------------
C
      NPRINT = 100
      DCURVE = 0.D0
      DPRESS = 0.D0
      DELPI = 0.D0
      DO 200 I=1,NSTEPS
      DO 100 J=1,1000
      DO 26 L=1,5
      X(L) = X0(L)+2.D0*STEP*Y(L,6)
   26 CONTINUE
C
C ---------------- Determine the functions {Gi} ----------------
C
      X11 = X(1)
      X12 = X11*X11
      X13 = X12*X11
      X14 = X13*X11
      X15 = X14*X11
      DLNX11 = DLOG(X11)
      G1 = -3.D0*X14+4.D0*X12-1.D0+4.D0*X14*DLNX11
      G1 = 1.D0/G1
      G2 = X13-X11-X13*DLNX11
      G2 = G2*G1
      G3 = -2.D0/5.D0+1.D0/20.D0*X15+1.D0/4.D0*X11-1.D0/6.D0*X13
      G3 = -0.25D0*(1.D0-X12)**2-G3/(1.D0-X11)
      G4 = -2.D0/5.D0+1.D0/20.D0*X15+1.D0/4.D0*X11-1.D0/6.D0*X13
      G4 = TW/DELT*(-X13+X11)-(-X13+X11)*X(2)+G4*X(2)/(1.D0-X11)**2
      G4 = G4-(1.D0-X12)**2/(1.D0-X11)*X(2)
      G5 = 1.D0/6.D0*X13*DLNX11-1.D0/9.D0+1.D0/4.D0*X11-5.D0/36.D0*X13
      G5 = -1.D0/2.D0*X12*DLNX11-1.D0/4.D0*(1.D0-X12)-G5/(1.D0-X11)
      G6 = 1.D0/6.D0*X13*DLNX11-1.D0/9.D0+1.D0/4.D0*X11-5.D0/36.D0*X13
      G6 = -G6*X(2)/(1.D0-X11)**2-TW/DELT*X11*DLNX11
      G6 = G6+X11*X(2)*DLNX11
      DUMMY = 1.D0/2.D0*X12*DLNX11+1.D0/4.D0*(1.D0-X12)
      G6 = G6+DUMMY*X(2)/(1.D0-X11)
      G7 = TW/(TW-TV)-X(2)
      G7 = DSQRT(G7)
C
C --------- Determine the matrix coefficients and constants -----------
C
      IF(Y(1,6).GT.TPOINT) THEN
      NFLAG = 0
      Y(1,1) = 1.D0
      Y(1,2) = 0.D0
      Y(1,3) = 0.D0
      Y(1,4) = 0.D0
      Y(1,5) = 0.D0
      Y(1,7) = X(5)
      Y(2,1) = 0.D0
      Y(2,2) = PI(6)*G2
      Y(2,3) = 0.D0
      Y(2,4) = -PI(1)
      Y(2,5) = 0.D0
      Y(2,7) = -PI(2)*G1*X(3)
      Y(2,4) = Y(2,4)/Y(2,2)
```

```
      Y(2,7) = Y(2,7)/Y(2,2)
      Y(2,2) = 1.D0
      Y(3,1) = 0.D0
      Y(3,2) = 2.D0*PIE*X(1)*X(5)
      Y(3,3) = 1.D0
      Y(3,4) = 0.D0
      Y(3,5) = 0.D0
      Y(3,7) = -X(1)*X(2)/(1.D0-X(1))
      Y(4,1) = 0.D0
      C7 = 2.D0*X(1)*X(1)*G5-G3
      C8 = (G4-2.D0*X(1)*X(1)*G6)*X(5)
      C9 = -X(1)*G5*PI(6)
      C10 = X(1)*X(1)*X(5)*G6*PI(6)
      C11 = PI(3)*X(2)
      Y(4,2) = C7*YIT(4)+2.D0*C9*YIT(2)+C10
      Y(4,3) = 0.D0
      Y(4,4) = C7*YIT(2)+C8
      Y(4,5) = 0.D0
      Y(4,7) = C7*YIT(2)*YIT(4)+C9*YIT(2)*YIT(2)-C11
      Y(4,2) = Y(4,2)/Y(4,4)
      Y(4,7) = Y(4,7)/Y(4,4)
      Y(4,4) = 1.D0
      Y(5,1) = 0.D0
      Y(5,2) = 0.D0
      DUMMY = (1.D0+X(5)*X(5))**1.5
      X5 = X(5)
      THETA = DATAN(-X5)
      Y(5,3) = -DUMMY*PI(4)*G7*DCOS(THETA)/SIGBAR/X(1)
      Y(5,4) = 0.D0
      Y(5,5) = 1.D0
      Y(5,7) = DUMMY*(4.D0*PI(1)/SIGBAR*(X(4)-1.D0)+PI(5)/SIGBAR/(1.D
     *0-X(1))**3)
      Y(5,7) = Y(5,7)+(1.D0+X(5))/X(1)
      ELSE
      NFLAG = NFLAG+1
      IF(NFLAG.EQ.1) THEN
      PRINT27
      WRITE(5,27)
   27 FORMAT('********* TRANSITION HAS OCCURRED *********')
      X05 = 1.D0+X0(5)*X0(5)
      X0(5) = 1.D0/DSQRT(X05)*X0(1)
      Y0(5) = 2.D0*FN*X0(1)*Y0(1)
      X5 = 1.D0+X(5)*X(5)
      X(5) = 1.D0/DSQRT(X5)*X(1)
      Y(5,6) = 2.D0*FN1*X(1)*Y(1,6)
      ENDIF
      Y(1,1) = 1.D0
      Y(1,2) = 0.D0
      Y(1,3) = 0.D0
      Y(1,4) = 0.D0
      Y(1,5) = 0.D0
      DUMMY = X(1)*X(1)/X(5)/X(5)-1.D0
      Y(1,7) = -DSQRT(DUMMY)
      Y(2,1) = 0.D0
      Y(2,2) = PI(6)*G2
      Y(2,3) = 0.D0
      Y(2,4) = -PI(1)
      Y(2,5) = 0.D0
      Y(2,7) = -PI(2)*G1*X(3)
      Y(2,4) = Y(2,4)/Y(2,2)
      Y(2,7) = Y(2,7)/Y(2,2)
      Y(2,2) = 1.D0
      Y(3,1) = 0.D0
      Y(3,2) = 2.D0*PIE*X(1)*Y(1,7)
      Y(3,3) = 1.D0
      Y(3,4) = 0.D0
      Y(3,5) = 0.D0
      Y(3,7) = -X(1)*X(2)/(1.D0-X(1))
      Y(4,1) = 0.D0
      C7 = 2.D0*X(1)*X(1)*G5-G3
      C8 = (G4-2.D0*X(1)*X(1)*G6)*Y(1,7)
      C9 = -X(1)*G5*PI(6)
      C10 = X(1)*X(1)*Y(1,7)*G6*PI(6)
      C11 = PI(3)*X(2)
      Y(4,2) = C7*YIT(4)+2.D0*C9*YIT(2)+C10
      Y(4,3) = 0.D0
      Y(4,4) = C7*YIT(2)+C8
      Y(4,5) = 0.D0
      Y(4,7) = C7*YIT(2)*YIT(4)+C9*YIT(2)*YIT(2)-C11
      Y(4,2) = Y(4,2)/Y(4,4)
      Y(4,7) = Y(4,7)/Y(4,4)
```

48-80

```fortran
      Y(4,4) = 1.D0
      Y(5,1) = 0.D0
      Y(5,2) = 0.D0
      D1 = Y(1,7)
      THETA = DATAN(-D1)
      Y(5,3) = Y(1,7)*PI(4)*G7*DCOS(THETA)/SIGBAR
      Y(5,4) = 0.D0
      Y(5,5) = 1.D0
      Y(5,7) = -X(1)*Y(1,7)*(4.D0*PI(1)/SIGBAR*(X(4)-1.D0)+PI(5)/SIGB
     *AR/(1.D0-X(1))**3)
      ENDIF
      DO 28 L=1,5
      XIT(L) = X(L)
      YIT(L) = Y(L,6)
   28 CONTINUE
C
C ------ Determine the solution vector using Gaussian elimination ------
C
      CALL GAUSSE(Y,N,NN)
C
C ------------------ Error criteria ------------------
C
      DO 29 K=1,5
      ERRY = (Y(K,6)-YIT(K))/Y(K,6)
      ERRY = DABS(ERRY)
      IF(ERRY.GT.ERROR)GO TO 100
      ERRX = (X(K)-XIT(K))/X(K)
      ERRX = DABS(ERRX)
      IF(ERRX.GT.ERROR)GO TO 100
   29 CONTINUE
      DUMMY = Y0(1)-Y(1,6)
      IF(NFLAG.GT.0.AND.DUMMY.LT.0.D0)GO TO 39
      GO TO 31
  100 CONTINUE
      PRINT30
   30 FORMAT('****** NO ERROR CONVERGENCE AFTER 1000 ITERATIONS ******')
      IF(J.GT.999)GO TO 1000
   31 XLOC = XLOC+STEP
C
C ------ Determine dimensionless curvature using the trapezoidal rule ------
C
C DCURVE - dimensionless mean interfacial curvature.
C
      IF(NFLAG.EQ.0) THEN
      FN = 1.D0+X0(5)*X0(5)
      FN = 1.D0/2.D0/X0(1)/DSQRT(FN)-1.D0/2.D0*Y0(5)/FN**(3.D0/2.D0)
      FN1 = 1.D0+X(5)*X(5)
      FN1 = 1.D0/2.D0/X(1)/DSQRT(FN1)-1.D0/2.D0*Y(5,6)/FN1**(3.D0/2.D
     *0)
      ENDIF
      IF(NFLAG.GT.0) THEN
      FN = Y0(5)/Y0(1)/X0(1)/2.D0
      FN1 = Y(5,6)/Y(1,6)/X(1)/2.D0
      ENDIF
      DCURVE = DCURVE+STEP/2.D0*(FN+FN1)
C
C --- Determine dimensionless pressure gradient using the trapezoidal rule ---
C
C DPRESS - dimensionless effective pressure difference.
C
      DPRESS = DPRESS + STEP/2.D0*(Y0(4)+Y(4,6))
C
C ------ Determine the change in pressure across the interface DELPI ------
C
C
C DELPI = dimensionless mean change in pressure across the interface.
C
      DPI0 = 1.D0/2.D0/PI(1)*SIGBAR*FN
      DPI0 = DPI0+PI(5)/4.D0/PI(1)/(1.D0-X0(1))**3
      DPI = 1.D0/2.D0/PI(1)*SIGBAR*FN1
      DPI = DPI+PI(5)/4.D0/PI(1)/(1.D0-X(1))**3
      DELPI = DELPI+STEP/2.D0*(DPI0+DPI)
C
C ------------------ Print results ------------------
C
      IF(NFLAG.GT.0.AND.X(5).LT.3.D-02)GO TO 39
      IF(I.LT.NPRINT)GO TO 37
      NPRINT = NPRINT+100
      WRITE(5,32)XLOC
      PRINT32,XLOC
   32 FORMAT(4X,'X = ',1PE13.6)
```

48-81

```
      WRITE(*,33)FN1
      WRITE(5,33)FN1
33    FORMAT(4X,'K = ',1PE13.6)
      WRITE(5,34)DPI
34    FORMAT(4X,'Local change in interfacial pressure = ',1PE13.6)
      DO 36 L=1,5
      WRITE(5,35)L,X(L),L,Y(L,6)
      PRINT35,L,X(L),L,Y(L,6)
35    FORMAT(2X,'X',I1,' = ',1PD21.14,6X,'Y',I1,' = ',1PD21.14)
36    CONTINUE
37    DO 38 II=1,5
      X0(II) = X(II)
      Y0(II) = Y(II,6)
38    CONTINUE
200   CONTINUE
39    WRITE(5,40)XLOC
40    FORMAT(4X,'XFINAL = ',1PE13.6)
      WRITE (5,41)FN1
41    FORMAT(4X,'K = ',1PE13.6)
      DO 43 L=1,5
      WRITE(5,42)L,X(L),L,Y(L,6)
42    FORMAT(2X,'X',I1,' = ',1PD21.14,6X,'Y',I1,' = ',1PD21.14)
43    CONTINUE
      XLOC0 = 1.D0-FN1*FN1*X(1)*X(1)
      IF(XLOC0.LT.0.) THEN
        WRITE(*,51)
        GO TO 1000
      ENDIF
      XLOC0 = XLOC+(1.D0-DSQRT(XLOC0))/FN1
      WRITE(5,44)XLOC0
44    FORMAT(4X,'Dimensionless X0 location = ',1PD13.6)
      DCURVE = (DCURVE+(XLOC0-XLOC)*FN1)/XLOC0
      WRITE(5,45)DCURVE
45    FORMAT(4X,'Dimensionless mean curvature = ',1PD13.6)
      DELPI = (DELPI+(XLOC0-XLOC)*DPI)/XLOC0
      WRITE(5,46)DELPI
46    FORMAT(4X,'Dimensionless mean interfacial pressure change = ',1PD1
     *3.6)
      WRITE(5,47)DPRESS
47    FORMAT(4X,'Dimensionless mean liquid pressure difference = ',1PD13
     *.6)
      WRITE(5,48)
48    FORMAT('Matrix coefficients')
      DO 50 K=1,7
      DO 50 L=1,5
      WRITE(5,49)L,K,Y(L,K)
49    FORMAT(2X,'Y',I1,I1,2X,' = ',1X,1PE13.6)
50    CONTINUE
51    FORMAT(2X,'******** The meniscus does not cross the x-axis *******
     **')
1000  STOP
      END
```

9. Appendix C

This appendix summarizes the results first derived by Lifshitz [1955] for the van der Waals interaction free energy of two half spaces filled with different media separated by a third medium.

The van der Waals free energy of interaction is given by

$$F(l,T) = \frac{K_B T}{8\pi l^2} \sum_{n=0}^{\infty} {}' I(i\xi_n, l),$$ (C.1)

where

$$I(i\xi_n, l) = \left(\frac{2\xi_n l\sqrt{\epsilon_2}}{c}\right)^2 \int_1^{\infty} dp\, p \ln D(i\xi_n; p),$$ (C.2)

In eqn. (C.2), the dispersion equation $D(i\xi_n; p)$ is defined as follows:

$$D(i\xi_n; p) = \left[1 - \overline{\Delta}_{12}\overline{\Delta}_{23} e^{-2\xi_n l p\sqrt{\epsilon_2}/c}\right]\left[1 - \Delta_{12}\Delta_{23} e^{-2\xi_n p\sqrt{\epsilon_2}/c}\right].$$ (C.3)

$$\overline{\Delta}_{i2} = \frac{s_i\epsilon_2 - p\epsilon_i}{s_i\epsilon_2 + p\epsilon_i} \quad , \qquad \Delta_{i2} = \frac{s_i - p}{s_i + p},$$ (C.4)

$$s_i = \sqrt{p^2 - 1 + \epsilon_i/\epsilon_2},$$ (C.5)

The sum in eqn. (C.1) is taken over imaginary frequencies given by $i\xi_n = 2\pi n K_B T/\hbar$, the prime on the sum denoting that the n = 0 term is multiplied by 1/2.

10. Appendix D

This appendix contains a complete listing of the Turbo Pascal program used to obtain the results given in section 4.3.2.

```
PROGRAM CapTube;

{$N+}

USES Crt,Dos;

CONST
   e       = 1.6021892e-19;
   c       = 2.99792458e8;
   KB      = 1.38066  .-23;
   hbar    = 1.0545887e-34;
   MaxN    = 500;
   MaxPts  = 300;
   SumTol  = 1.0e-5;

TYPE
   DataArray = ARRAY [0..1,0..MaxPts] OF DOUBLE;

VAR
   OutfileName              : STRING[12];
   n,mat,liq,count,step     : INTEGER;
   alpha,beta,Zn,Zs,pmax    : DOUBLE;
   E1,E2,E3,T,l,rc,ri,Eint  : DOUBLE;
   Fconst,Term,Sum,rat      : DOUBLE;
   Spect                    : DataArray;
   W,X                      : ARRAY [1..16] OF DOUBLE;
   Results                  : ARRAY [1..20,1..2] OF DOUBLE;

LABEL
   ExitNLoop;

(*****************************************
*                                       *
*  Section I : General Purpose Routines  *
*                                       *
*****************************************)

(----------------------------------------------------------+
|                                                          |
|   Generate Appropriate Error Message and Terminate       |
|                                                          |
+----------------------------------------------------------)
```

```
PROCEDURE ErrorHandler( Error : INTEGER );
BEGIN
  WriteLn;
  WriteLn('*** ERROR ***');
  CASE Error OF
    1  : WriteLn(' - Input file exceeds allocated memory');
    2  : WriteLn(' - Invalid material selection');
    3  : WriteLn(' - Invalid liquid selection');
    4  : WriteLn(' - Temperature out of range');
    5  : WriteLn(' - Capillary radius out of range');
    6  : WriteLn(' - Unable to open data file');
    7  : WriteLn(' - Specified tolerance not attained in
sum');
  END; { case }
  Halt;
END; { Procedure ErrorHandler }

{------------------------------------+
 |                                    |
 |    Load Optical Spectrum of Metal  |
 |                                    |
 +------------------------------------}

PROCEDURE LoadSpectrum( opt : INTEGER; VAR Spect :
DataArray;
                           VAR count : INTEGER );
VAR
  i               : INTEGER;
  w,Nopt,Kopt : DOUBLE;
  InfileName  : STRING[12];
  Infile      : TEXT;
BEGIN
  CASE opt OF
    1 : BEGIN ( Load gold data )
          InfileName := 'Gold.opt';
          alpha      :=  6.12298;
          beta       := -2.60488;
        END;
    2 : BEGIN ( Load Copper data )
          InfileName := 'Copper.opt';
          alpha      :=  5.70688;
          beta       := -2.51289;
        END;
    3 : BEGIN ( Load Silver data )
          InfileName := 'Silver.opt';
          alpha      :=  4.88767;
          beta       := -2.70051;
        END;
  END; ( Case )
  Assign(Infile,InfileName);
  ReSet(Infile);
```

```
      IF IOResult <> 0 THEN ErrorHandler(6);
      count := 0;
      FOR i := 0 TO MaxPts DO BEGIN
        Spect[0,i] := 0.0;
        Spect[1,i] := 0.0;
      END;
      WHILE NOT Eof(Infile) DO BEGIN
         Read(Infile,w,Nopt,Kopt);
        Spect[0,count] := w;
        Spect[1,count] := Nopt*Kopt;
        Inc(count);
        IF count > MaxPts THEN ErrorHandler(1);
      END; { while }
      Close(Infile);
END; { Procedure LoadSpectrum }

{-------------------------------------------------+
 |                                                 |
 |    Display Menu and Choose Materials for Run    |
 |                                                 |
 +-------------------------------------------------}

PROCEDURE DisplayMenu;
BEGIN { DisplayMenu Main }
  ClrScr;
  WriteLn;
  WriteLn('+---------------------------------------------+
');
  WriteLn('|                                             |
');
  WriteLn('|   Hamaker Constants in Metal Capillary Tubes  |
');
  WriteLn('|                                             |
');
  WriteLn('+---------------------------------------------+
');
  WriteLn;
  WriteLn('Material Options        Liquid Options ');
  WriteLn('----------------        -------------- ');
  WriteLn;
  WriteLn('1 - Gold              1 - Pentane ');
  WriteLn('2 - Copper            2 - Hexane  ');
  WriteLn('3 - Silver            3 - Heptane ');
  WriteLn('4 - Platinum          4 - Octane  ');
  WriteLn;
  Write('Material?            > ');
  ReadLn(mat);
  IF mat > 4 THEN ErrorHandler(2);
  LoadSpectrum(mat,Spect,count);
  Write('Liquid?              > ');
  ReadLn(liq);
```

```pascal
    IF liq > 4 THEN ErrorHandler(3);
    Write('Temperature?        > ');
    ReadLn(T);
    IF (T < 50) OR (T > 400) THEN ErrorHandler(4);
    Write('Layer thickness?    > ');
    ReadLn(l);
    Write('Maxp?               > ');
    ReadLn(pmax);
    Write('Output file name? > ');
    ReadLn(OutfileName);
    WriteLn;
    WriteLn('Running...');
    WriteLn;
END; { Procedure DisplayMenu }


(*************************************************
 *                                               *
 *  Section II : Dielectric Properties Routines  *
 *                                               *
 *************************************************)


{---------------------------------------------}
{-                                           -}
{-  Low Frequency Extrapolation Function     -}
{-                                           -}
{---------------------------------------------}

FUNCTION Extrapol( omega,xi : DOUBLE ) : DOUBLE;
BEGIN
    Extrapol := Exp((beta-1))*ln(xi)/(Sqr(omega)+Sqr(xi));
END; { Function Extrapol }


{-------------------------------------------------+
 |                                                 |
 |    Composite Simpson Integration Routine        |
 |                                                 |
 +-------------------------------------------------}

FUNCTION CSimpson( lower,upper,xi : DOUBLE;
     NumStep : INTEGER ) : DOUBLE;
VAR
   omega,h,EndSum,EvenSum,OddSum : DOUBLE;
   i                             : INTEGER;
BEGIN
   h         := (upper-lower)/(2*NumStep);
   EndSum    := Extrapol(lower,xi)+Extrapol(upper,xi);
   EvenSum   := 0.0;
   OddSum    := 0.0;
   FOR i := 1 TO 2*NumStep-1 DO BEGIN
     omega := lower+i*h;
     IF Odd(i) THEN OddSum   := OddSum+Extrapol(omega,xi)
```

```
                ELSE EvenSum := EvenSum+Extrapol(omega,xi);
  END; { for }
  CSimpson := h*(FndSum+2*EvenSum+4*OddSum)/3;
END; { Function CSimpson }


{--------------------------------------------------+
 |                                                  |
 |    Use Kramers-Kronig Relation to Calculate E1   |
 |                                                  |
 +--------------------------------------------------}


FUNCTION Kramers( xi : DOUBLE ) : DOUBLE;
VAR
  i                        : INTEGER;
  ksum,h,lowterm,E21,E22 : DOUBLE;
BEGIN { Kramers Main }
  xi    := hbar*xi/e; { Convert to eV }
  ksum := 0.0;
  FOR i := 0 TO count-2 DO BEGIN
    h     := Spect[0,i+1]-Spect[0,i];
     E21  := Spect[0,i]*Spect[1,i]/
    (Sqr(Spect[0,i])+Sqr(xi));
     E22  := Spect[0,i+1]*Spect[1,i+1]/
    (Sqr(Spect[0,i+1])+Sqr(xi));
    ksum := ksum+h*(E21+E22);
  END; { for }
  lowterm := CSimpson(0.01,Spect[0,0],xi,50);
  Kramers := 1+(2/Pi)*(alpha*lowterm+ksum);
END; { Function Kramers }


{------------------------------------------+
 |                                          |
 |    Calculate Dielectric Permitivities    |
 |                                          |
 +------------------------------------------}


PROCEDURE CalculateE( xi : DOUBLE; VAR E1,E2,E3 : DOUBLE );
BEGIN
  E1 := Kramers(xi);
  E2 := 1.00;
  CASE liq OF
    1 : E3 := 1+0.025/(1+Sqr(xi/5.540e14))
     +0.819/(1+Sqr(xi/1.877e16));
    2 : E3 := 1+0.026/(1+Sqr(xi/5.540e14))
     +0.864/(1+Sqr(xi/1.873e16));
    3 : E3 := 1+0.025/(1+Sqr(xi/5.540e14))
     +0.898/(1+Sqr(xi/1.870e16));
    4 : E3 := 1+0.023/(1+Sqr(xi/5.540e14))
     +0.925/(1+Sqr(xi/1.863e16));
  END; { liq case }
END; { Procedure CalculateE }
```

```
(************************************************************
*                                                          *
*   Section III : van der Waals Interaction Calculations   *
*                                                          *
************************************************************)

(----------------------------------------------------+
|                                                     |
|       Dispersion Relation for a Capillary Tube      |
|                                                     |
+----------------------------------------------------)

FUNCTION Dtube( pz,pt : DOUBLE ) : DOUBLE;
VAR
  delta11,delta12,delta21,delta22 : EXTENDED;
  p,p1,p2,p3,psqr,corr,tmp,tmp2   : EXTENDED;
BEGIN
  psqr    := (Sqr(pz)+Sqr(pt));
  p1      := Sqrt(psqr+E1/E3);
  p2      := Sqrt(psqr+E2/E3);
  p3      := Sqrt(psqr+1);
  delta11 := (p1*E3-E1*p3)/(p1*E3+p3*E1);
  delta12 := (p2*E3-p3*E2)/(p2*E3+p3*E2);
  delta21 := (p1-p3)/(p1+p3);
  delta22 := (p2-p3)/(p2+p3);
  corr    := 1+rat*(1/rc)*pt;
  tmp     := corr*Exp(-rat*p3);
  tmp2    :=
(1-delta11*delta12*tmp)*(1-delta21*delta22*tmp);
  Dtube   := ln(tmp2);
END; ( Function Dtube )

(----------------------------------------------------+
|                                                     |
|      Initialize Weights and Abscissas for Gauss     |
|                                                     |
+----------------------------------------------------)

PROCEDURE InitGauss;
BEGIN
  X[1]  := 0.048307665687738; W[1]  := 0.096540088514727;
  X[2]  := 0.144471961582796; W[2]  := 0.095638720079275;
  X[3]  := 0.239287362252137; W[3]  := 0.093844399080804;
  X[4]  := 0.331868602282128; W[4]  := 0.091173878695764;
  X[5]  := 0.421351276130635; W[5]  := 0.087652093004404;
  X[6]  := 0.506899908932229; W[6]  := 0.083311924226947;
  X[7]  := 0.587715757240762; W[7]  := 0.078193895787070;
  X[8]  := 0.663044266930215; W[8]  := 0.072345794108848;
  X[9]  := 0.732182118740290; W[9]  := 0.065822222776362;
  X[10] := 0.794483795967942; W[10] := 0.058684093478535;
  X[11] := 0.849367613732570; W[11] := 0.050998059262376;
```

```
    X[12] := 0.896321155766052; W[12] := 0.042835898022227;
    X[13] := 0.934906075937739; W[13] := 0.034273862913021;
    X[14] := 0.964762255587506; W[14] := 0.025392065309262;
    X[15] := 0.985611511545268; W[15] := 0.016274394730906;
    X[16] := 0.997263861849481; W[16] := 0.007018610009470;
END; ( Procedure InitGauss )

{----------------------------------------+
|                                         |
|   32 Point Gauss-Legendre Integration   |
|                                         |
+----------------------------------------}

FUNCTION Gauss2D( minpz,maxpz,minpt,maxpt : DOUBLE ) :
DOUBLE;
VAR
    i,j                       : INTEGER;
    pzr,pzm,dpz,ptm,ptr,dpt  : EXTENDED;
    gterm,gsum,tmp1,tmp2     : EXTENDED;
BEGIN
    pzm := 0.5*(maxpz+minpz);
    pzr := 0.5*(maxpz-minpz);
    ptm := 0.5*(maxpt+minpt);
    ptr := 0.5*(maxpt-minpt);
    gsum := 0.0;
    FOR i := 1 TO 16 DO BEGIN
        dpz    := pzr*X[i];
        gterm := 0.0;
        FOR j := 1 TO 16 DO BEGIN
      dpt    := ptr*X[j];
      tmp1   := Dtube(pzr+dpz,ptr+dpt)+Dtube(pzr-dpz,ptr+dpt);
      tmp2   := Dtube(pzr+dpz,ptr-dpt)+Dtube(pzr-dpz,ptr-dpt);
      gterm := gterm+W[j]*(tmp1+tmp2);
        END; ( for j )
        gsum := gsum+W[i]*gterm;
    END; ( for i )
    Gauss2D := pzr*ptr*gsum;
END; ( Function Gauss2D )

{----------------------------------+
|                                   |
|   Display Summary of Results      |
|                                   |
+----------------------------------}

PROCEDURE WriteSummary;
VAR
    OutFile    : TEXT;
    index      : INTEGER;
BEGIN
    Assign(Outfile,OutfileName);
```

```
   ReWrite(Outfile);
   WriteLn(Outfile,1);
   CASE liq OF
      1 : WriteLn(Outfile,'Pentane');
      2 : WriteLn(Outfile,'Hexane');
      3 : WriteLn(Outfile,'Heptane');
      4 : WriteLn(Outfile,'Octane');
   END; { case }
   CASE mat OF
      1 : WriteLn(Outfile,'Gold');
      2 : WriteLn(Outfile,'Copper');
      3 : WriteLn(Outfile,'Silver');
   END; { case }
   FOR index := 1 TO 12 DO BEGIN
      WriteLn(Outfile,Results[index,1],' ',Results[index,2]);
   END; { for index }
   Close(Outfile);
END; { Procedure WriteSummary }

BEGIN { Program CapTube Main }
   WriteLn;
   DisplayMenu;
   WriteLn;
   Fconst := 2*Pi*KB*T/hbar; { Frequency constant }
   InitGauss;
   FOR step := 1 TO 12 DO BEGIN
      rc   := step*5.00e-6;
      ri   := rc-l;
      IF l > 0.1*rc THEN ErrorHandler(5);
      sum := 0.0;
      FOR n := 1 TO MaxN DO BEGIN { n loop }
     Zn := Fconst*n;
     CalculateE(Zn,E1,E2,E3);
     Zs    := c/(2*l*sqrt(E3));
     rat   := Zn/Zs;
     Term := Sqr(rat)*Gauss2D(0.0,pmax,0.0,pmax);
     Sum   := Sum+Term;
     Eint := (KB*T*Sum)/Sqr(2*Pi*l);
     IF Abs(Term/Sum) < SumTol THEN GOTO ExitNLoop;
      END; { for n }
     ErrorHandler(7);
      ExitNLoop:
      Results[step,1] := rc;
      Results[step,2] := Eint;
      WriteLn('  ',Round(rc/1.00e-6) : 5,'  ',Eint);
   END; { for step }
   Sound(220);
   Delay(500);
   NoSound;
```

```
    WriteLn;
    WriteSummary;
END. { Program CapTube }
```

11. Appendix E

This appendix is a complete listing of the data files gold.opt and silver.opt which are used by the program listed in Appendix D.

listing of gold.opt:

```
0.125 12.24 54.7 0.130 11.51 53.1 0.135 10.84 51.6
0.140 10.21 50.2 0.145 9.582 48.8 0.150 9.016 47.6
0.155 8.500 46.4 0.160 8.060 45.2 0.165 7.655 44.1
0.170 7.282 43.0 0.175 6.937 42.0 0.180 6.598 41.0
0.185 6.270 40.1 0.190 5.966 39.1 0.195 5.684 38.3
0.200 5.423 37.5 0.210 4.971 35.9 0.220 4.611 34.3
0.230 4.292 32.9 0.240 4.007 31.7 0.250 3.748 30.5
0.260 3.507 29.3 0.270 3.289 28.2 0.280 3.091 27.2
0.290 2.912 26.3 0.300 2.749 25.4 0.310 2.600 24.6
0.320 2.471 23.9 0.330 2.352 23.1 0.340 2.242 22.5
0.350 2.141 21.9 0.360 2.046 21.3 0.370 1.958 20.7
0.380 1.876 20.2 0.390 1.800 19.7 0.400 1.728 19.2
0.420 1.598 18.3 0.440 1.482 17.5 0.460 1.379 16.8
0.480 1.287 16.1 0.500 1.205 15.5 0.520 1.130 14.9
0.540 1.063 14.4 0.560 1.002 13.9 0.580 0.947 13.4
0.600 0.896 13.0 0.620 0.850 12.6 0.640 0.807 12.2
0.660 0.767 11.9 0.680 0.730 11.5 0.700 0.696 11.2
0.720 0.665 10.9 0.740 0.636 10.6 0.760 0.609 10.3
0.780 0.583 10.1 0.800 0.559 9.81 0.820 0.537 9.58
0.840 0.515 9.36 0.860 0.493 9.15 0.880 0.473 8.96
0.900 0.454 8.77 0.920 0.436 8.59 0.940 0.419 8.42
0.960 0.403 8.25 0.980 0.389 8.09 1.000 0.372 8.77
1.100 0.312 7.93 1.200 0.272 7.07 1.300 0.236 6.47
1.400 0.210 5.88 1.500 0.188 5.39 1.600 0.174 4.86
1.700 0.164 4.35 1.800 0.160 3.80 1.900 0.166 3.15
2.200 0.306 2.88 2.300 0.402 2.540 2.400 0.608 2.120
2.500 0.916 1.840 2.600 1.252 1.796 2.700 1.426 1.846
2.800 1.562 1.904 2.900 1.616 1.940 3.000 1.636 1.958
3.100 1.658 1.956 3.200 1.674 1.936 3.300 1.696 1.906
3.400 1.716 1.862 3.500 1.740 1.848 3.600 1.766 1.846
3.700 1.798 1.860 3.800 1.824 1.878 3.900 1.840 1.904
4.000 1.830 1.916 4.100 1.812 1.920 4.200 1.776 1.918
4.300 1.742 1.900 4.400 1.690 1.882 4.500 1.648 1.852
4.600 1.598 1.822 4.700 1.546 1.784 4.800 1.504 1.748
4.900 1.490 1.698 5.000 1.484 1.636 5.100 1.478 1.590
5.200 1.470 1.550 5.300 1.462 1.510 5.400 1.454 1.478
5.500 1.452 1.442 5.600 1.442 1.418 5.700 1.438 1.388
5.800 1.432 1.364 5.900 1.430 1.334 6.000 1.422 1.306
6.199 1.427 1.215 6.358 1.424 1.170 6.526 1.427 1.135
6.702 1.442 1.107 6.888 1.470 1.085 7.085 1.500 1.070
7.293 1.519 1.070 7.514 1.512 1.093 7.749 1.483 1.106
```

```
7.999 1.450 1.108 8.266 1.419 1.102 8.551 1.386 1.089
8.856 1.360 1.072 8.920 1.355 1.067 8.948 1.350 1.063
9.050 1.345 1.058 9.117 1.338 1.053 9.184 1.333 1.050
9.253 1.328 1.045 9.322 1.323 1.038 9.393 1.318 1.032
9.465 1.313 1.027 9.537 1.308 1.020 9.611 1.304 1.017
9.686 1.300 1.012 9.763 1.295 1.005 9.840 1.290 1.000
9.919 1.285 0.994 10.00 1.280 0.987 10.08 1.275 0.982
10.16 1.270 0.975 10.25 1.265 0.967 10.33 1.260 0.962
10.42 1.255 0.955 10.51 1.250 0.950 10.60 1.247 0.942
10.69 1.242 0.935 10.78 1.237 0.932 10.88 1.232 0.927
10.97 1.228 0.922 11.07 1.225 0.914 11.17 1.223 0.907
11.27 1.222 0.900 11.37 1.220 0.893 11.48 1.218 0.885
11.59 1.218 0.878 11.70 1.217 0.870 11.81 1.217 0.865
11.92 1.215 0.860 12.04 1.213 0.853 12.16 1.210 0.848
12.28 1.207 0.842 12.40 1.200 0.836 12.52 1.190 0.831
12.65 1.180 0.826 12.78 1.170 0.823 12.92 1.159 0.819
13.05 1.146 0.815 13.19 1.133 0.812 13.33 1.121 0.809
13.48 1.110 0.806 13.62 1.100 0.801 13.78 1.090 0.798
13.93 1.080 0.793 14.09 1.070 0.789 14.25 1.061 0.784
14.42 1.053 0.780 14.59 1.048 0.775 14.76 1.041 0.770
14.94 1.037 0.765 15.12 1.033 0.759 15.31 1.030 0.752
15.50 1.029 0.745 15.69 1.028 0.739 15.90 1.028 0.730
16.10 1.029 0.720 16.31 1.030 0.713 16.53 1.033 0.704
16.75 1.038 0.697 16.98 1.042 0.690 17.22 1.050 0.685
17.46 1.060 0.680 17.71 1.075 0.678 17.97 1.088 0.680
18.23 1.107 0.687 18.51 1.125 0.694 18.79 1.140 0.700
19.07 1.155 0.710 19.37 1.157 0.730 19.68 1.153 0.750
20.00 1.146 0.770 20.33 1.134 0.795 20.66 1.113 0.813
21.01 1.085 0.830 21.38 1.055 0.842 21.75 1.022 0.850
22.14 0.985 0.848 22.54 0.950 0.840 22.96 0.915 0.825
23.39 0.890 0.795 23.84 0.872 0.740 24.31 0.860 0.695
24.80 0.850 0.645 25.30 0.846 0.600 25.83 0.846 0.565
26.00 0.896 0.679 26.38 0.855 0.548 27.00 0.895 0.635
28.00 0.901 0.607 29.00 0.905 0.599 29.50 0.900 0.601
30.00 0.887 0.603 30.50 0.870 0.601 31.00 0.851 0.596
32.00 0.805 0.570 33.00 0.781 0.515 34.00 0.779 0.473
35.00 0.783 0.439 36.00 0.786 0.414 37.00 0.788 0.391
38.00 0.791 0.371 39.00 0.793 0.353 40.00 0.795 0.334
41.00 0.799 0.317 42.00 0.804 0.303 43.00 0.807 0.292
44.00 0.807 0.282 45.00 0.805 0.270 46.00 0.805 0.254
47.00 0.809 0.238 48.00 0.814 0.225 49.00 0.821 0.213
50.00 0.828 0.204 52.00 0.839 0.190 54.00 0.848 0.180
56.00 0.856 0.171 58.00 0.861 0.166 60.00 0.863 0.162
62.00 0.859 0.156 64.00 0.857 0.147 66.00 0.856 0.136
68.00 0.855 0.124 70.00 0.857 0.112 72.00 0.862 9.970e-2
74.00 0.869 9.270e-2 76.00 0.873 8.710e-2 78.00 0.875
8.180e-2
80.00 0.878 7.520e-2 82.00 0.881 6.820e-2 83.00 0.884
6.530e-2
83.50 0.885 6.370e-2 84.00 0.888 6.330e-2 84.50 0.888
6.540e-2
```

```
85.00 0.887 6.480e-2 86.00 0.886 6.130e-2 87.00 0.887
5.760e-2
87.50 0.888 5.580e-2

listing of silver.opt:

0.125 13.11 53.70 0.130 12.21 52.20 0.140 10.69 49.40
0.150 9.441 47.10 0.160 8.376 44.80 0.170 7.461 42.50
0.180 6.670 40.40 0.190 5.960 38.60 0.200 5.355 37.00
0.220 4.425 34.00 0.240 3.732 31.30 0.260 3.202 29.00
0.280 2.786 26.90 0.300 2.446 25.10 0.320 2.160 23.50
0.340 1.915 22.10 0.360 1.710 20.90 0.380 1.536 19.80
0.400 1.387 18.80 0.420 1.265 17.90 0.440 1.168 17.10
0.460 1.083 16.40 0.480 1.007 15.70 0.500 0.939 15.10
0.520 0.878 14.50 0.540 0.823 14.00 0.560 0.774 13.50
0.580 0.729 13.00 0.600 0 568 12.60 0.620 0.650 12.20
0.640 0.624 11.80 0.650 0.873 13.30 0.660 0.600 11.40
0.680 0.578 11.10 0.700 0.557 10.70 0.720 0.537 10.40
0.740 0.519 10.10 0.750 0.624 11.50 0.760 0.501 9.840
0.780 0.485 9.570 0.800 0.469 9.320 0.820 0.455 9.080
0.840 0.442 8.880 0.850 0.446 10.10 0.860 0.431 8.700
0.880 0.421 8.370 0.900 0.411 8.370 0.920 0.401 8.210
0.940 0.392 8.060 0.950 0.358 8.950 0.960 0.383 7.920
0.980 0.375 7.780 1.000 0.329 8.490 1.100 0.251 7.670
1.200 0.226 6.990 1.300 0.198 6.430 1.400 0.163 5.950
1.500 0.145 5.500 1.600 0.143 5.090 1.700 0.148 4.740
1.800 0.140 4.440 1.900 0.140 4.150 2.000 0.131 3.880
2.100 0.121 3.660 2.200 0.120 3.450 2.300 0.129 3.250
2.400 0.130 3.070 2.500 0.130 2.880 2.600 0.132 2.720
2.700 0.144 2.560 2.800 0.157 2.400 2.900 0.160 2.260
3.000 0.173 2.110 3.100 0.173 1.950 3.200 0.192 1.810
3.300 0.200 1.670 3.400 0.186 1.610 3.500 0.209 1.440
3.600 0.238 1.240 3.650 0.259 1.120 3.700 0.294 0.986
3.730 0.321 0.902 3.750 0.371 0.813 3.800 0.526 0.663
3.830 0.616 0.609 3.850 0.708 0.565 3.880 0.815 0.526
3.900 0.932 0.504 3.930 1.044 0.514 3.950 1.149 0.540
3.980 1.247 0.586 4.000 1.323 0.647 4.050 1.432 0.766
4.100 1.496 0.882 4.150 1.522 0.992 4.200 1.519 1.080
4.300 1.502 1.190 4.400 1.476 1.260 4.500 1.441 1.310
4.600 1.404 1.330 4.700 1.372 1.350 4.800 1.343 1.350
4.900 1.320 1.350 5.000 1.298 1.350 5.200 1.265 1.330
5.400 1.238 1.310 5.600 1.208 1.300 5.800 1.173 1.290
6.000 1.125 1.270 6.100 1.098 1.260 6.200 1.072 1.240
6.300 1.048 1.210 6.400 1.028 1.180 6.600 0.995 1.130
6.800 0.969 1.070 7.000 0.953 1.010 7.200 0.942 0.951
7.400 0.936 0.892 7.600 0.935 0.832 7.800 0.940 0.770
8.000 0.962 0.706 8.200 0.993 0.653 8.400 1.032 0.610
8.600 1.073 0.581 8.800 1.112 0.563 9.000 1.149 0.552
9.200 1.182 0.550 9.700 1.229 0.566 10.00 1.241 0.568
10.30 1.252 0.564 10.60 1.265 0.560 10.90 1.280 0.560
11.20 1.293 0.566 11.50 1.300 0.573 11.80 1.304 0.578
```

```
12.10 1.308 0.581 12.50 1.315 0.588 13.00 1.322 0.603
13.50 1.322 0.628 14.00 1.308 0.656 14.50 1.281 0.679
15.00 1.243 0.693 15.50 1.202 0.691 16.00 1.167 0.676
17.00 1.121 0.635 18.00 1.101 0.593 19.00 1.096 0.565
20.00 1.098 0.548 20.50 1.105 0.549 21.00 1.109 0.561
21.50 1.104 0.584 22.00 1.075 0.615 22.50 1.029 0.624
23.00 0.980 0.614 23.50 0.943 0.584 24.50 0.906 0.522
25.50 0.891 0.468 26.00 0.887 0.444 26.50 0.886 0.421
27.00 0.886 0.650 27.50 0.851 0.616 28.00 0.902 0.590
28.50 0.911 0.572 29.00 0.919 0.557 30.00 0.931 0.541
31.00 0.932 0.534 32.00 0.921 0.528 33.00 0.899 0.514
34.00 0.879 0.489 35.00 0.865 0.454 36.00 0.876 0.418
37.00 0.885 0.400 38.00 0.892 0.388 39.00 0.895 0.378
40.00 0.896 0.368 42.00 0.897 0.349 44.00 0.899 0.334
46.00 0.896 0.323 47.00 0.893 0.317 48.00 0.888 0.309
49.00 0.885 0.300 50.00 0.884 0.290 51.00 0.886 0.281
52.00 0.889 0.276 53.00 0.890 0.275 54.00 0.885 0.274
56.00 0.871 0.259 58.00 0.868 0.237 60.00 0.873 0.221
62.00 0.881 0.211 64.00 0.885 0.210 66.00 0.883 0.211
68.00 0.871 0.208 70.00 0.859 0.197 72.00 0.851 0.181
74.00 0.853 0.166 76.00 0.853 0.158 78.00 0.851 0.149
80.00 0.848 0.139 85.00 0.846 0.111 90.00 0.849 8.24e-2
95.00 0.861 5.65e-2 100.0 0.876 3.82e-2 105.0 0.890 2.61e-2
110.0 0.902 1.72e-2 120.0 0.926 6.22e-3 130.0 0.943 3.66e-3
140.0 0.955 3.54e-3 150.0 0.963 4.07e-3 160.0 0.968 4.38e-3
170.0 0.973 4.42e-3 180.0 0.978 4.31e-3 200.0 0.982 3.77e-3
220.0 0.987 3.45e-3 240.0 0.989 3.50e-3 260.0 0.992 3.11e-3
280.0 0.994 2.70e-3 300.0 0.996 2.32e-3 350.0 1.001 1.68e-3
370.0 1.007 8.23e-3 400.0 0.998 8.82e-3 500.0 0.997 4.57e-3
600.0 0.997 2.69e-3 740.0 0.998 1.42e-3 800.0 0.997 1.08e-3
1000 0.998 5.23e-4 1500 0.999 1.21e-4 2000 1.001 4.72e-5
2500 1.001 2.40e-5 3000 1.001 1.43e-5 3500 1.003 1.72e-5
3600 1.003 3.24e-5 4000 1.003 3.57e-5 4500 1.001 2.21e-5
5000 1.001 1.52e-5 6000 1.002 8.36e-6 8000 1.000 2.92e-6
10000 0.999 1.29e-6
```

12. Appendix F

Table F.1: Results for Pentane on Gold, $l = 10nm$.

$r_c(\mu)$	$F_{tube}(J/m^2)$	$\delta_{surf}(\%)$	$\delta_{Ham}(\%)$
5	1.309002E-05	2.860E-01	2.987E-01
10	1.307136E-05	1.430E-01	1.497E-01
15	1.306514E-05	9.536E-02	9.986E-02
20	1.306203E-05	7.153E-02	7.492E-02
25	1.306016E-05	5.723E-02	5.995E-02
30	1.305892E-05	4.770E-02	4.997E-02
35	1.305803E-05	4.089E-02	4.283E-02
40	1.305736E-05	3.578E-02	3.748E-02
45	1.305684E-05	3.181E-02	3.332E-02
50	1.305643E-05	2.863E-02	2.999E-02
55	1.305609E-05	2.603E-02	2.726E-02
60	1.305581E-05	2.387E-02	2.499E-02

$F_{plate} = 1.305269E-05 \ J/m^2$

Table F.2: Results for Pentane on Gold, $l = 30nm$.

$r_c(\mu)$	$F_{tube}(J/m^2)$	$\delta_{surf}(\%)$	$\delta_{Ham}(\%)$
5	7.930640E-07	8.990E-01	8.887E-01
10	7.895310E-07	4.495E-01	4.472E-01
15	7.883533E-07	2.997E-01	2.987E-01
20	7.877644E-07	2.248E-01	2.243E-01
25	7.874111E-07	1.798E-01	1.795E-01
30	7.871755E-07	1.498E-01	1.497E-01
35	7.870073E-07	1.284E-01	1.283E-01
40	7.868811E-07	1.124E-01	1.123E-01
45	7.867829E-07	9.989E-02	9.986E-02
50	7.867044E-07	8.990E-02	8.989E-02
55	7.866402E-07	8.173E-02	8.172E-02
60	7.865866E-07	7.492E-02	7.492E-02

$F_{plate} = 7.859978E-07 \ J/m^2$

Table F.3: Results for Pentane on Gold, $l = 50nm$.

$r_c(\mu)$	$F_{tube}(J/m^2)$	$\delta_{surf}(\%)$	$\delta_{Ham}(\%)$
5	1.981788E-07	1.513E+00	1.469E+00
10	1.967015E-07	7.568E-01	7.421E-01
15	1.962091E-07	5.045E-01	4.965E-01
20	1.959628E-07	3.784E-01	3.730E-01
25	1.958151E-07	3.027E-01	2.987E-01
30	1.957166E-07	2.523E-01	2.491E-01
35	1.956462E-07	2.162E-01	2.136E-01
40	1.955935E-07	1.892E-01	1.870E-01
45	1.955524E-07	1.682E-01	1.663E-01
50	1.955196E-07	1.514E-01	1.497E-01
55	1.954927E-07	1.376E-01	1.361E-01
60	1.954703E-07	1.261E-01	1.248E-01

$F_{plate} = 1.952241E-07 \ J/m^2$

Table F.4: Results for Pentane on Gold, $l = 70nm$.

$r_c(\mu)$	$F_{tube}(J/m^2)$	$\delta_{surf}(\%)$	$\delta_{Ham}(\%)$
5	7.794090E-08	2.129E+00	2.038E+00
10	7.712827E-08	1.064E+00	1.035E+00
15	7.685759E-08	7.094E-01	6.931E-01
20	7.672225E-08	5.321E-01	5.211E-01
25	7.664105E-08	4.256E-01	4.175E-01
30	7.658691E-08	3.547E-01	3.483E-01
35	7.654824E-08	3.040E-01	2.987E-01
40	7.651924E-08	2.660E-01	2.615E-01
45	7.649668E-08	2.365E-01	2.326E-01
50	7.647863E-08	2.128E-01	2.094E-01
55	7.646387E-08	1.935E-01	1.904E-01
60	7.645156E-08	1.774E-01	1.746E-01

$F_{plate} = 7.631621E-08 \ J/m^2$

Table F.5: Results for Pentane on Gold, $l = 90nm$.

$r_c(\mu)$	$F_{tube}(J/m^2)$	$\delta_{surf}(\%)$	$\delta_{Ham}(\%)$
5	3.847558E-08	2.744E+00	2.598E+00
10	3.796166E-08	1.372E+00	1.325E+00
15	3.779045E-08	9.145E-01	8.887E-01
20	3.770484E-08	6.859E-01	6.686E-01
25	3.765347E-08	5.487E-01	5.359E-01
30	3.761922E-08	4.573E-01	4.472E-01
35	3.759476E-08	3.919E-01	3.836E-01
40	3.757642E-08	3.429E-01	3.359E-01
45	3.756215E-08	3.048E-01	2.987E-01
50	3.755073E-08	2.744E-01	2.690E-01
55	3.754139E-08	2.494E-01	2.446E-01
60	3.753361E-08	2.286E-01	2.243E-01

$F_{plate} = 3.744799E-08 \; J/m^2$

Table F.6: Results for Pentane on Silver, $l = 10nm$.

$r_c(\mu)$	$F_{tube}(J/m^2)$	$\delta_{surf}(\%)$	$\delta_{Ham}(\%)$
5	1.127879E-05	2.848E-01	2.987E-01
10	1.126277E-05	1.424E-01	1.497E-01
15	1.125743E-05	9.490E-02	9.986E-02
20	1.125476E-05	7.116E-02	7.492E-02
25	1.125316E-05	5.692E-02	5.995E-02
30	1.125209E-05	4.743E-02	4.997E-02
35	1.125133E-05	4.065E-02	4.283E-02
40	1.125076E-05	3.556E-02	3.748E-02
45	1.125031E-05	3.161E-02	3.332E-02
50	1.124996E-05	2.844E-02	2.999E-02
55	1.124967E-05	2.585E-02	2.726E-02
60	1.124943E-05	2.370E-02	2.499E-02

$F_{plate} = 1.124676E-05 \; J/m^2$

Table F.7: Results for Pentane on Silver, $l = 30 nm$.

$r_c (\mu)$	$F_{tube} (J/m^2)$	$\delta_{surf} (\%)$	$\delta_{Ham} (\%)$
5	7.080257E-07	8.918E-01	8.887E-01
10	7.048967E-07	4.459E-01	4.472E-01
15	7.038536E-07	2.973E-01	2.987E-01
20	7.033321E-07	2.230E-01	2.243E-01
25	7.030192E-07	1.784E-01	1.795E-01
30	7.028106E-07	1.486E-01	1.497E-01
35	7.026616E-07	1.274E-01	1.283E-01
40	7.025498E-07	1.115E-01	1.123E-01
45	7.024629E-07	9.909E-02	9.986E-02
50	7.023933E-07	8.918E-02	8.989E-02
55	7.023364E-07	8.107E-02	8.172E-02
60	7.022890E-07	7.432E-02	7.492E-02

$F_{plate} = 7.017675E-07 \; J/m^2$

Table F.8: Results for Pentane on Silver, $l = 50 nm$.

$r_c (\mu)$	$F_{tube} (J/m^2)$	$\delta_{surf} (\%)$	$\delta_{Ham} (\%)$
5	1.820135E-07	1.503E+00	1.469E+00
10	1.806664E-07	7.513E-01	7.421E-01
15	1.802174E-07	5.009E-01	4.965E-01
20	1.799928E-07	3.757E-01	3.730E-01
25	1.798581E-07	3.005E-01	2.987E-01
30	1.797683E-07	2.504E-01	2.491E-01
35	1.797041E-07	2.147E-01	2.136E-01
40	1.796560E-07	1.878E-01	1.870E-01
45	1.796186E-07	1.670E-01	1.663E-01
50	1.795887E-07	1.503E-01	1.497E-01
55	1.795642E-07	1.366E-01	1.361E-01
60	1.795438E-07	1.252E-01	1.248E-01

$F_{plate} = 1.793192E-07 \; J/m^2$

Table F.9: Results for Pentane on Silver, $l = 70nm$.

$r_c(\mu)$	$F_{tube}(J/m^2)$	$\delta_{surf}(\%)$	$\delta_{Ham}(\%)$
5	7.295093E-08	2.115E+00	2.038E+00
10	7.219541E-08	1.058E+00	1.035E+00
15	7.194355E-08	7.052E-01	6.931E-01
20	7.181762E-08	5.289E-01	5.211E-01
25	7.174205E-08	4.231E-01	4.175E-01
30	7.169168E-08	3.526E-01	3.483E-01
35	7.165569E-08	3.022E-01	2.987E-01
40	7.162871E-08	2.644E-01	2.615E-01
45	7.160772E-08	2.351E-01	2.326E-01
50	7.159092E-08	2.116E-01	2.094E-01
55	7.157719E-08	1.923E-01	1.904E-01
60	7.156574E-08	1.763E-01	1.746E-01

$F_{plate} = 7.143979E{-}08 \ J/m^2$

Table F.10: Results for Pentane on Silver, $l = 90nm$.

$r_c(\mu)$	$F_{tube}(J/m^2)$	$\delta_{surf}(\%)$	$\delta_{Ham}(\%)$
5	3.648110E-08	2.730E+00	2.598E+00
10	3.599641E-08	1.365E+00	1.325E+00
15	3.583483E-08	9.101E-01	8.887E-01
20	3.575403E-08	6.826E-01	6.686E-01
25	3.570556E-08	5.461E-01	5.359E-01
30	3.567324E-08	4.551E-01	4.472E-01
35	3.565015E-08	3.901E-01	3.836E-01
40	3.563284E-08	3.413E-01	3.359E-01
45	3.561937E-08	3.034E-01	2.987E-01
50	3.560860E-08	2.731E-01	2.690E-01
55	3.559978E-08	2.482E-01	2.446E-01
60	3.559244E-08	2.276E-01	2.243E-01

$F_{plate} = 3.551163E{-}08 \ J/m^2$

Table F.11: Results for Octane on Gold, $l = 10nm$.

$r_c(\mu)$	$F_{tube}(J/m^2)$	$\delta_{surf}(\%)$	$\delta_{Ham}(\%)$
5	1.375089E-05	2.871E-01	2.987E-01
10	1.373128E-05	1.440E-01	1.497E-01
15	1.372474E-05	9.633E-02	9.986E-02
20	1.372133E-05	7.151E-02	7.492E-02
25	1.371937E-05	5.720E-02	5.995E-02
30	1.371807E-05	4.766E-02	4.997E-02
35	1.371713E-05	4.085E-02	4.283E-02
40	1.371643E-05	3.574E-02	3.748E-02
45	1.371589E-05	3.177E-02	3.332E-02
50	1.371545E-05	2.859E-02	2.999E-02
55	1.371509E-05	2.599E-02	2.726E-02
60	1.371480E-05	2.382E-02	2.499E-02

$F_{plate} = 1.371153E-05 \; J/m^2$

Table F.12: Results for Octane on Gold, $l = 30nm$.

$r_c(\mu)$	$F_{tube}(J/m^2)$	$\delta_{surf}(\%)$	$\delta_{Ham}(\%)$
5	8.315127E-07	8.988E-01	8.887E-01
10	8.278092E-07	4.494E-01	4.472E-01
15	8.265747E-00	2.996E-01	2.987E-01
20	8.259574E-07	2.247E-01	2.243E-01
25	8.255870E-07	1.798E-01	1.795E-01
30	8.253401E-07	1.498E-01	1.497E-01
35	8.251637E-07	1.284E-01	1.283E-01
40	8.250315E-07	1.124E-01	1.123E-01
45	8.249286E-07	9.988E-02	9.986E-02
50	8.248463E-07	8.989E-02	8.989E-02
55	8.247789E-07	8.172E-02	8.172E-02
60	8.247228E-07	7.491E-02	7.492E-02

$F_{plate} = 8.241055E-07 \; J/m^2$

Table F.13: Results for Octane on Gold, $l = 50nm$.

$r_c(\mu)$	$F_{tube}(J/m^2)$	$\delta_{surf}(\%)$	$\delta_{Ham}(\%)$
5	2.079322E-07	1.513E+00	1.469E+00
10	2.063828E-07	7.565E-01	7.421E-01
15	2.058663E-07	5.044E-01	4.965E-01
20	2.056080E-07	3.783E-01	3.730E-01
25	2.054531E-07	3.026E-01	2.987E-01
30	2.053497E-07	2.522E-01	2.491E-01
35	2.052760E-07	2.162E-01	2.136E-01
40	2.052206E-07	1.891E-01	1.870E-01
45	2.051776E-07	1.681E-01	1.663E-01
50	2.051431E-07	1.513E-01	1.497E-01
55	2.051150E-07	1.376E-01	1.361E-01
60	2.050915E-07	1.261E-01	1.248E-01

$F_{plate} = $ 2.048332E-07 J/m^2

Table F.14: Results for Octane on Gold, $l = 70nm$.

$r_c(\mu)$	$F_{tube}(J/m^2)$	$\delta_{surf}(\%)$	$\delta_{Ham}(\%)$
5	8.182831E-08	2.127E+00	2.038E+00
10	8.097609E-08	1.064E+00	1.035E+00
15	8.069200E-08	7.092E-01	6.931E-01
20	8.054994E-08	5.319E-01	5.211E-01
25	8.046471E-08	4.255E-01	4.175E-01
30	8.040788E-08	3.546E-01	3.483E-01
35	8.036730E-08	3.039E-01	2.987E-01
40	8.033685E-08	2.660E-01	2.615E-01
45	8.031318E-08	2.364E-01	2.326E-01
50	8.029424E-08	2.128E-01	2.094E-01
55	8.027874E-08	1.934E-01	1.904E-01
60	8.026582E-08	1.773E-01	1.746E-01

$F_{plate} = $ 8.012376E-08 J/m^2

Table F.15: Results for Octane on Gold, $l = 90nm$.

$r_c(\mu)$	$F_{tube}(J/m^2)$	$\delta_{surf}(\%)$	$\delta_{Ham}(\%)$
5	4.040896E-08	2.743E+00	2.598E+00
10	3.986965E-08	1.371E+00	1.325E+00
15	3.968986E-08	9.144E-01	8.887E-01
20	3.959996E-08	6.858E-01	6.686E-01
25	3.954601E-08	5.486E-01	5.359E-01
30	3.951005E-08	4.572E-01	4.472E-01
35	3.948437E-08	3.919E-01	3.836E-01
40	3.946510E-08	3.429E-01	3.359E-01
45	3.945012E-08	3.048E-01	2.987E-01
50	3.943813E-08	2.743E-01	2.690E-01
55	3.942832E-08	2.494E-01	2.446E-01
60	3.942015E-08	2.286E-01	2.243E-01

$F_{plate} = 3.933024E-08 \ J/m^2$

Table F.16: Results for Octane on Silver, $l = 10nm$.

$r_c(\mu)$	$F_{tube}(J/m^2)$	$\delta_{surf}(\%)$	$\delta_{Ham}(\%)$
5	1.175066E-05	2.848E-01	2.987E-01
10	1.173398E-05	1.424E-01	1.497E-01
15	1.172842E-05	9.496E-02	9.986E-02
20	1.172564E-05	7.123E-02	7.492E-02
25	1.172397E-05	5.699E-02	5.995E-02
30	1.172285E-05	4.749E-02	4.997E-02
35	1.172206E-05	4.071E-02	4.283E-02
40	1.172146E-05	3.563E-02	3.748E-02
45	1.172100E-05	3.167E-02	3.332E-02
50	1.172063E-05	2.851E-02	2.999E-02
55	1.172033E-05	2.592E-02	2.726E-02
60	1.172007E-05	2.376E-02	2.499E-02

$F_{plate} = 1.171729E-05 \ J/m^2$

Table F.17: Results for Octane on Silver, $l = 30 nm$.

$r_c(\mu)$	$F_{tube}(J/m^2)$	$\delta_{surf}(\%)$	$\delta_{Ham}(\%)$
5	7.391851E-07	8.908E-01	8.887E-01
10	7.359220E-07	4.454E-01	4.472E-01
15	7.348343E-07	2.969E-01	2.987E-01
20	7.342904E-07	2.227E-01	2.243E-01
25	7.339641E-07	1.782E-01	1.795E-01
30	7.337465E-07	1.485E-01	1.497E-01
35	7.335911E-07	1.273E-01	1.283E-01
40	7.334746E-07	1.113E-01	1.123E-01
45	7.333840E-07	9.898E-02	9.986E-02
50	7.333114E-07	8.908E-02	8.989E-02
55	7.332521E-07	8.098E-02	8.172E-02
60	7.332027E-07	7.423E-02	7.492E-02

$F_{plate} = 7.326588E-07 \ J/m^2$

Table F.18: Results for Octane on Silver, $l = 50 nm$.

$r_c(\mu)$	$F_{tube}(J/m^2)$	$\delta_{surf}(\%)$	$\delta_{Ham}(\%)$
5	1.905673E-07	1.501E+00	1.469E+00
10	1.891583E-07	7.505E-01	7.421E-01
15	1.886886E-07	5.004E-01	4.965E-01
20	1.884538E-07	3.753E-01	3.730E-01
25	1.883128E-07	3.002E-01	2.987E-01
30	1.882189E-07	2.502E-01	2.491E-01
35	1.881518E-07	2.144E-01	2.136E-01
40	1.881015E-07	1.876E-01	1.870E-01
45	1.880623E-07	1.668E-01	1.663E-01
50	1.880310E-07	1.501E-01	1.497E-01
55	1.880054E-07	1.364E-01	1.361E-01
60	1.879840E-07	1.251E-01	1.248E-01

$F_{plate} = 1.877492E-07 \ J/m^2$

Table F.19: Results for Octane on Silver, $l = 70nm$.

$r_c(\mu)$	$F_{tube}(J/m^2)$	$\delta_{surf}(\%)$	$\delta_{Ham}(\%)$
5	7.650618E-08	2.114E+00	2.038E+00
10	7.571440E-08	1.057E+00	1.035E+00
15	7.545044E-08	7.047E-01	6.931E-01
20	7.531846E-08	5.285E-01	5.211E-01
25	7.523927E-08	4.228E-01	4.175E-01
30	7.518647E-08	3.523E-01	3.483E-01
35	7.514876E-08	3.020E-01	2.987E-01
40	7.512048E-08	2.643E-01	2.615E-01
45	7.509848E-08	2.349E-01	2.326E-01
50	7.508088E-08	2.114E-01	2.094E-01
55	7.506649E-08	1.922E-01	1.904E-01
60	7.505449E-08	1.762E-01	1.746E-01

$F_{plate} = 7.492249E-08 \ J/m^2$

Table F.20: Results for Octane on Silver, $l = 90nm$.

| $r_c(\mu)$	$F_{tube}(J/m^2)$	$\delta_{surf}(\%)$	$\delta_{Ham}(\%)$
5	3.829272E-08	2.730E+00	2.598E+00
10	3.778421E-08	1.365E+00	1.325E+00
15	3.761469E-08	9.105E-01	8.887E-01
20	3.752992E-08	6.831E-01	6.686E-01
25	3.747906E-08	5.467E-01	5.359E-01
30	3.744515E-08	4.557E-01	4.472E-01
35	3.742093E-08	3.907E-01	3.836E-01
40	3.740276E-08	3.420E-01	3.359E-01
45	3.738864E-08	3.041E-01	2.987E-01
50	3.737733E-08	2.738E-01	2.690E-01
55	3.736808E-08	2.490E-01	2.446E-01
60	3.736038E-08	2.283E-01	2.243E-01

$F_{plate} = 3.727528E-08 \ J/m^2$
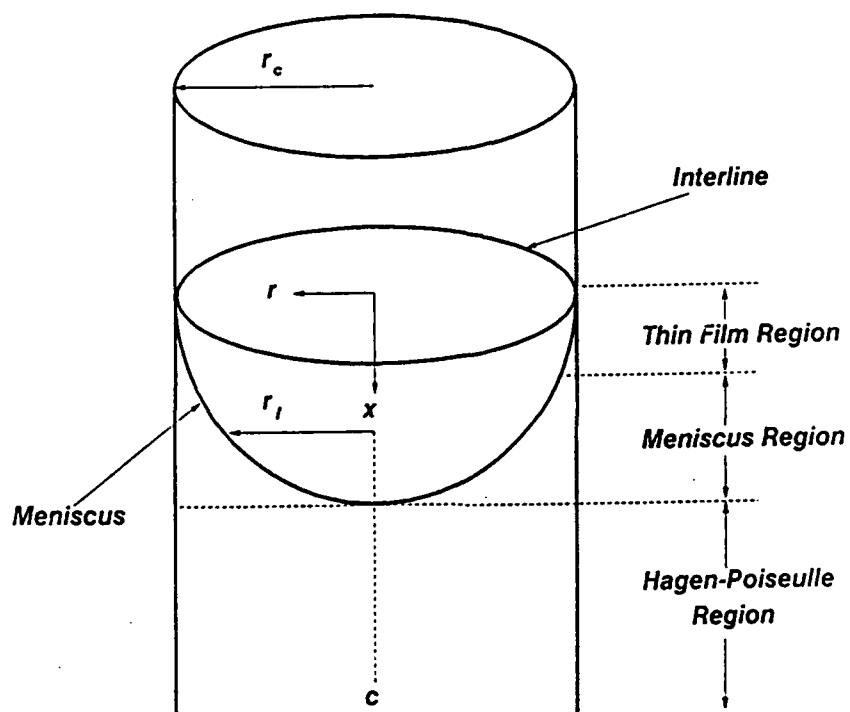
Figure 3.1. Flow Regions for an Isolated Evaporating Meniscus in a Capillary Tube.
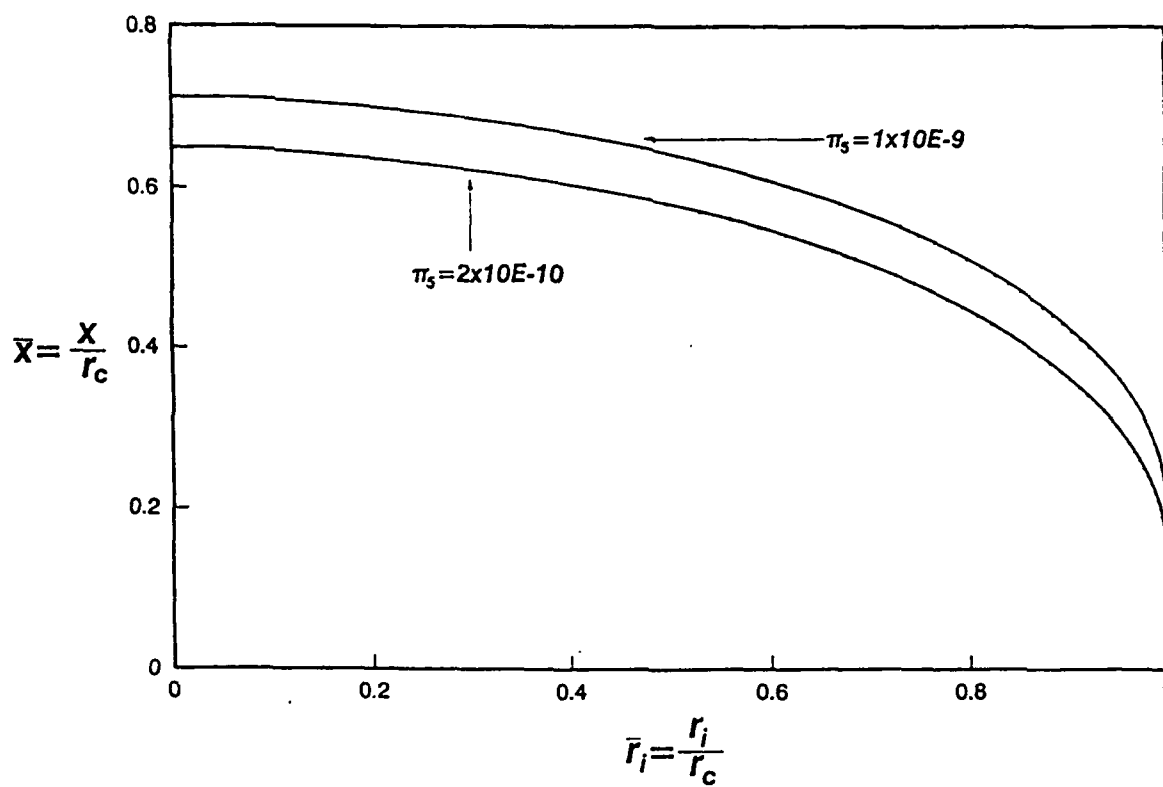
Figure 3.2.   Meniscus Profiles for Different Values of the Dispersion Number.
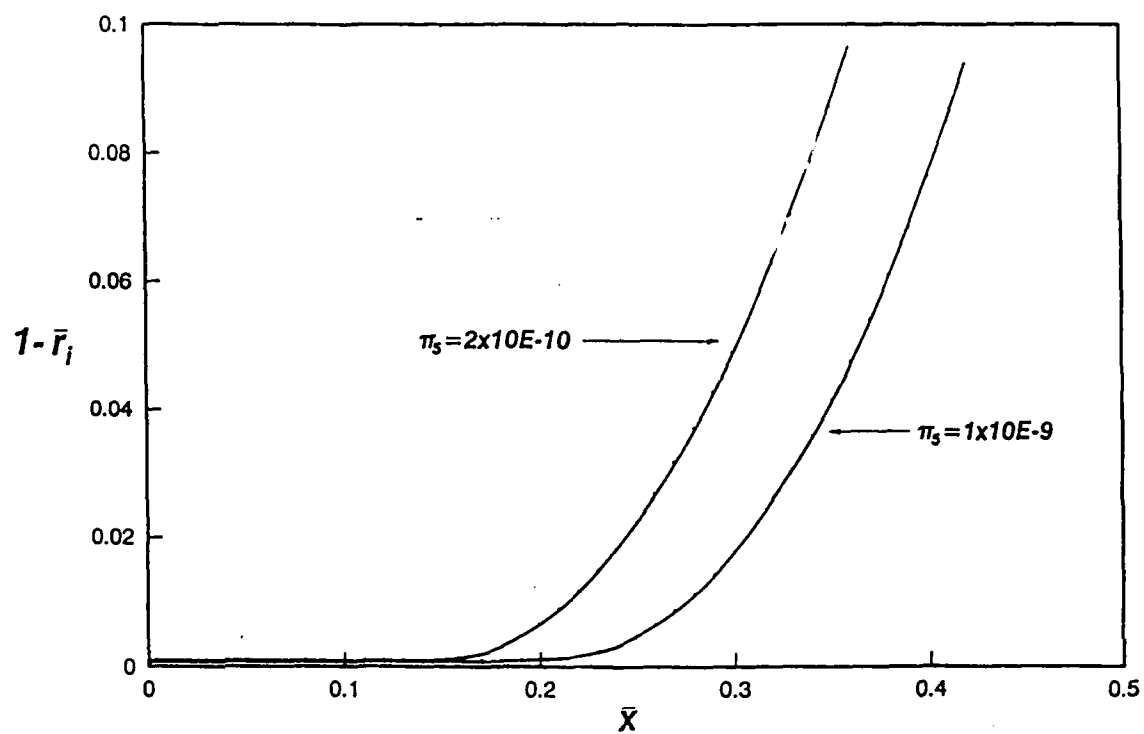
Figure 3.3.   Microscopic Meniscus Profiles for Different Values of the Dispersion Number.

$\pi_s = 2 \times 10E\text{-}10$

$\pi_s = 6 \times 10E\text{-}10$

$\pi_s = 1 \times 10E\text{-}9$
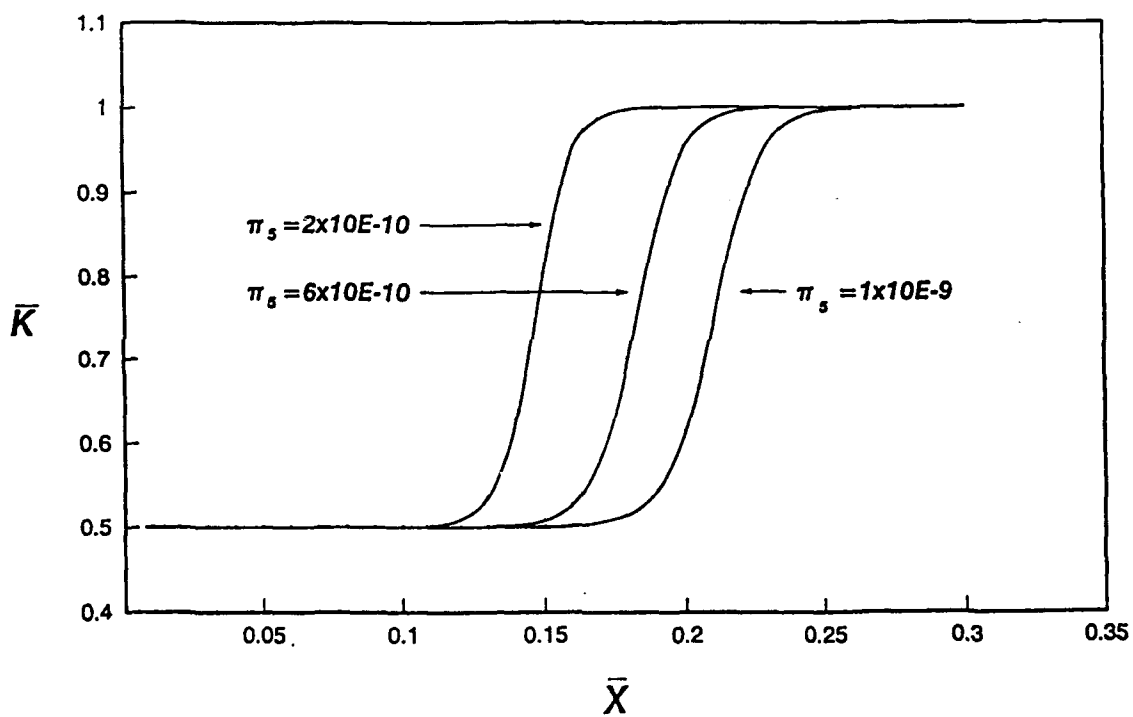
Figure   Dimensionless Mean Curvature Profile
3.4.   for Different Values of the Dispersion
         Number.

The graph shows:

- Y-axis label: $-\bar{m}''x10^6$
- Y-axis values: 0, 0.5, 1, 1.5, 2, 2.5
- X-axis label: $\bar{X}$
- X-axis values: 0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6

Equation shown on graph:
$$\bar{m}'' = \frac{r_c h_{fg}}{k_L(T_w - T_v)} m''$$

Curve labels:
- $\pi_2 = 2.47 \times 10E\text{-}11$
- $\pi_2 = 7.55 \times 10E\text{-}11$
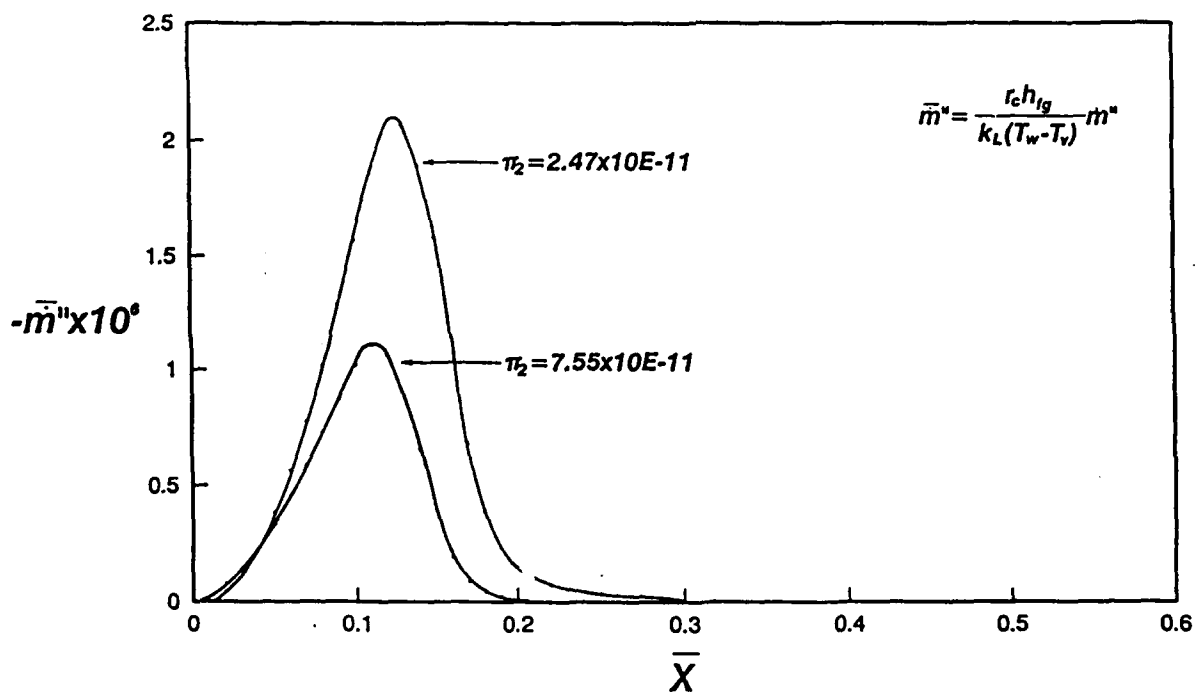
Figure 3.5. Dimensionless Interfacial Mass Flux Profile for $\pi_5 = 2 \times 10^{-10}$.
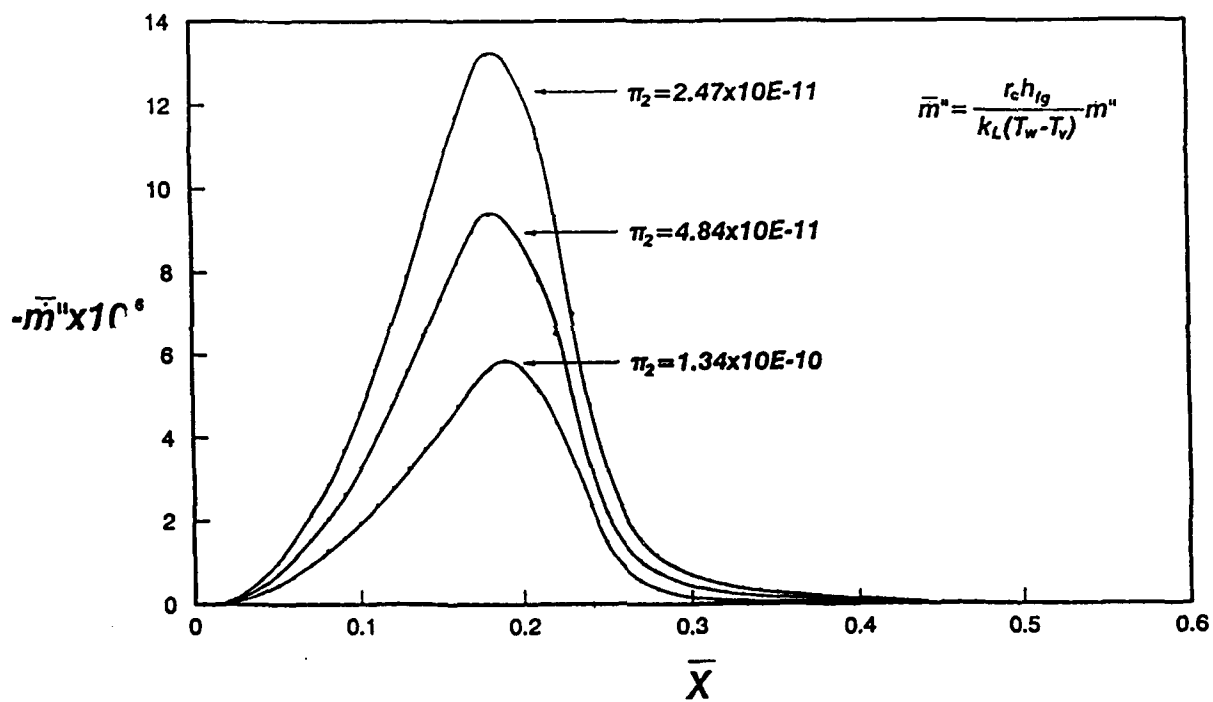
Figure 3.6. Dimensionless Interfacial Mass Flux Profile for $\pi_5 = 1 \times 10^{-9}$.
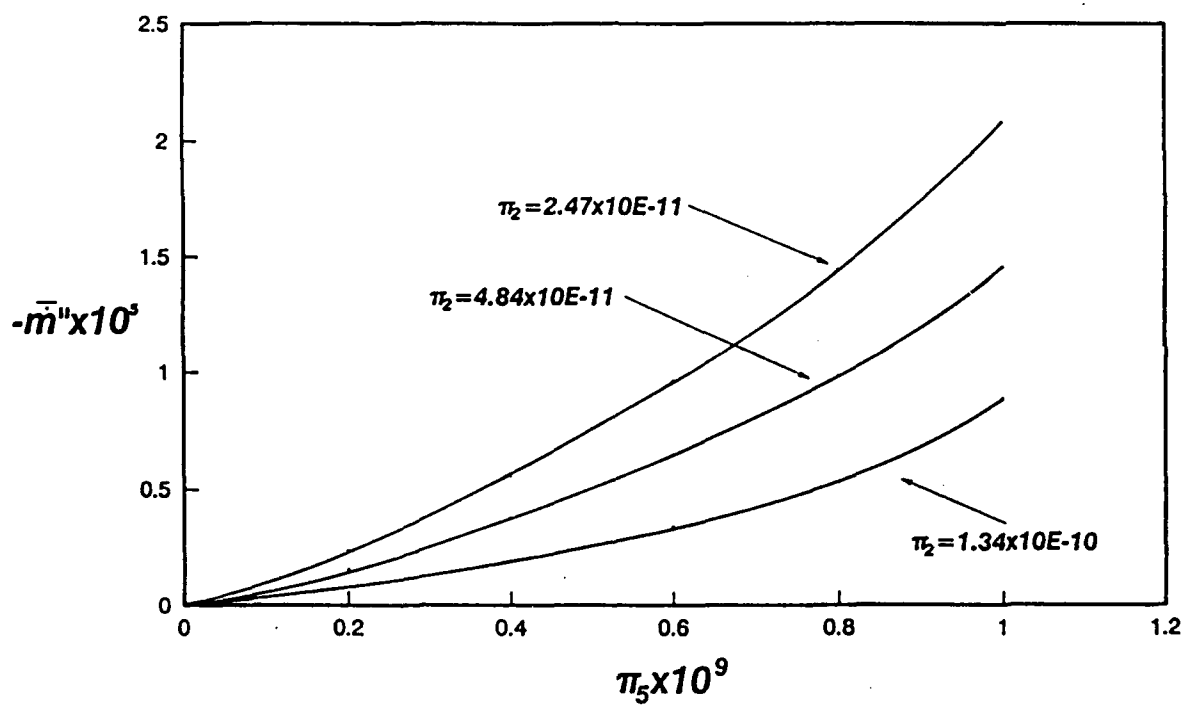
Figure 3.7. Dimensionless Mass Transfer Rate as a Function of the Dispersion Number.

$$\bar{P}_{vi} - \bar{P}_{ll}$$

All Computer Results and $\frac{2\sigma}{r_c}$

$r_c x 10^{-4} (m)$
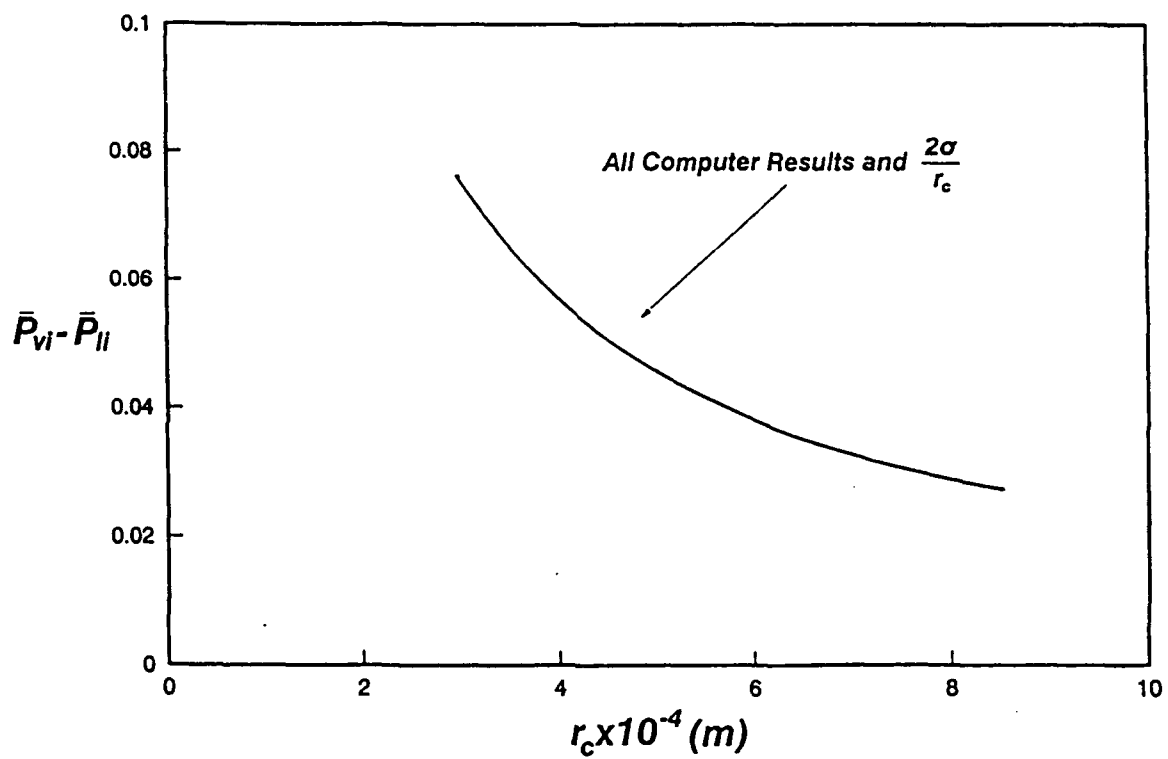
Figure 3.8. Dimensionless Capillary Pressure as a Function of the Capillary Radius for all values of π_2 and π_5.

log Fplate

———— Hamaker Theo

········· With Retardati

-5

-6

-7

100 20 30 40 50 60 70 80 90

Condensed Liquid Phase Thickness (nm)
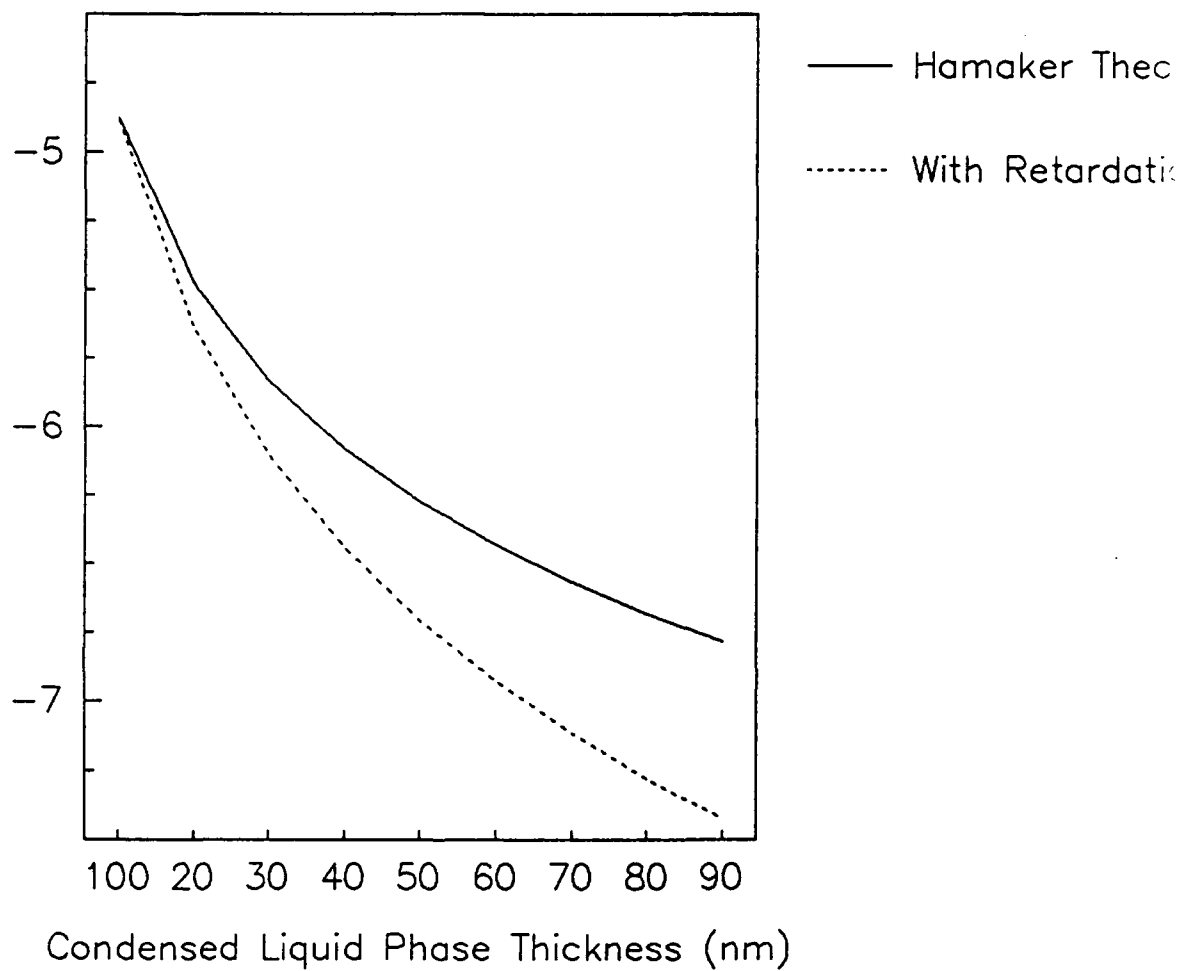
Figure 4.1. Comparision of the van der Waals interaction predicted by the Hamaker theory and surface mode analysis.

Percent Deviation
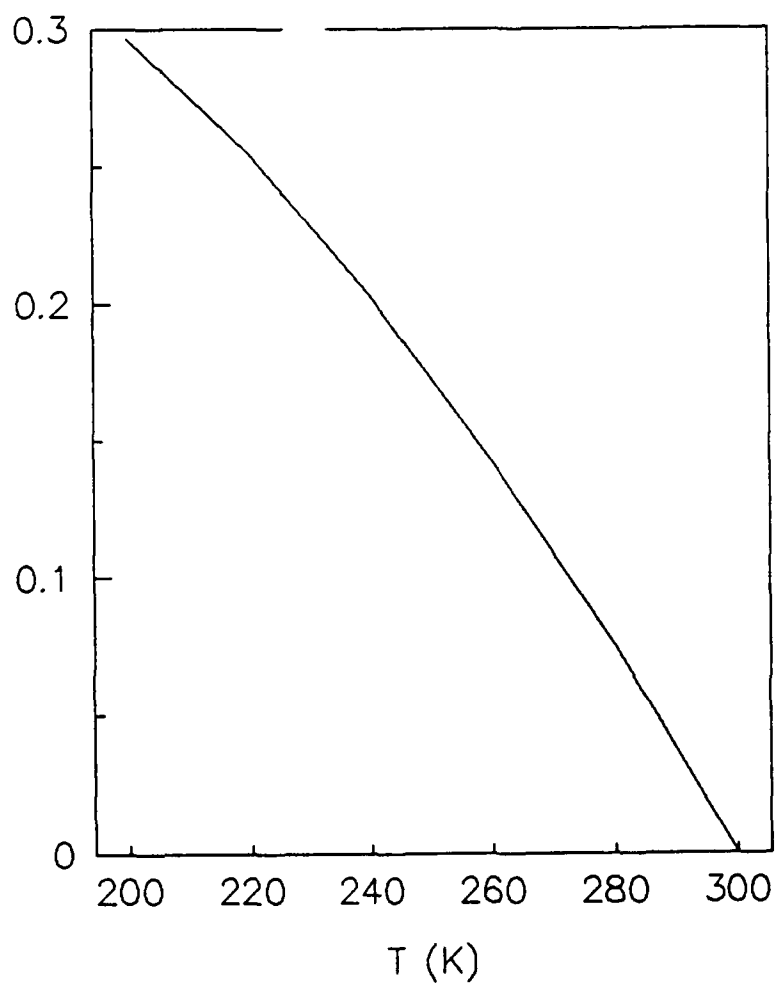
Figure 4.2. Percent deviation of van der Waals interaction from value at 300 K as a function of temperature.

Deviation (percent)

Capillary Radius (microns)
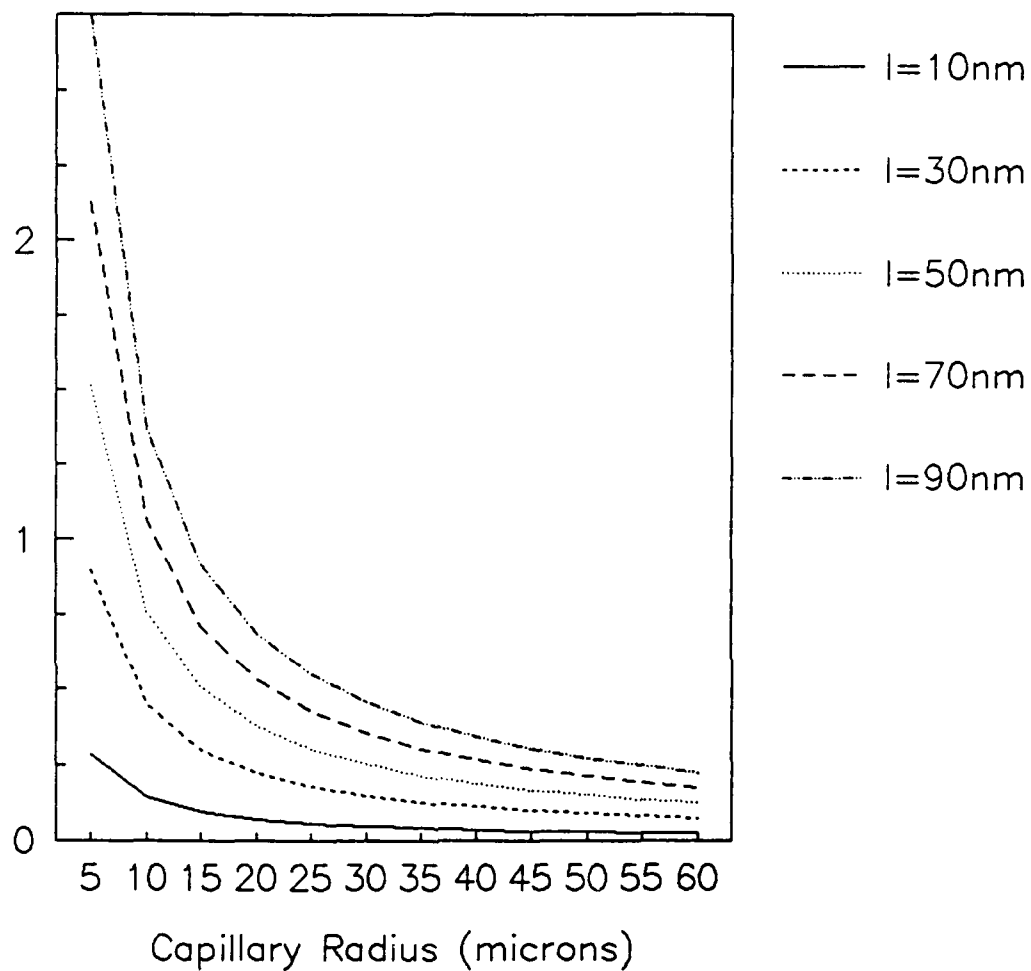
Figure 4.3. Percent deviation of the van der Waals interaction in a capillary tube from that in a flat plate geometry as a function of capillary radius.

Ratio of Percent Deviations

Capillary Radius (microns)

Legend:
- —— l=10nm
- ········ l=30nm
- ········ l=50nm
- - - - l=70nm
- —·—·— l=90nm
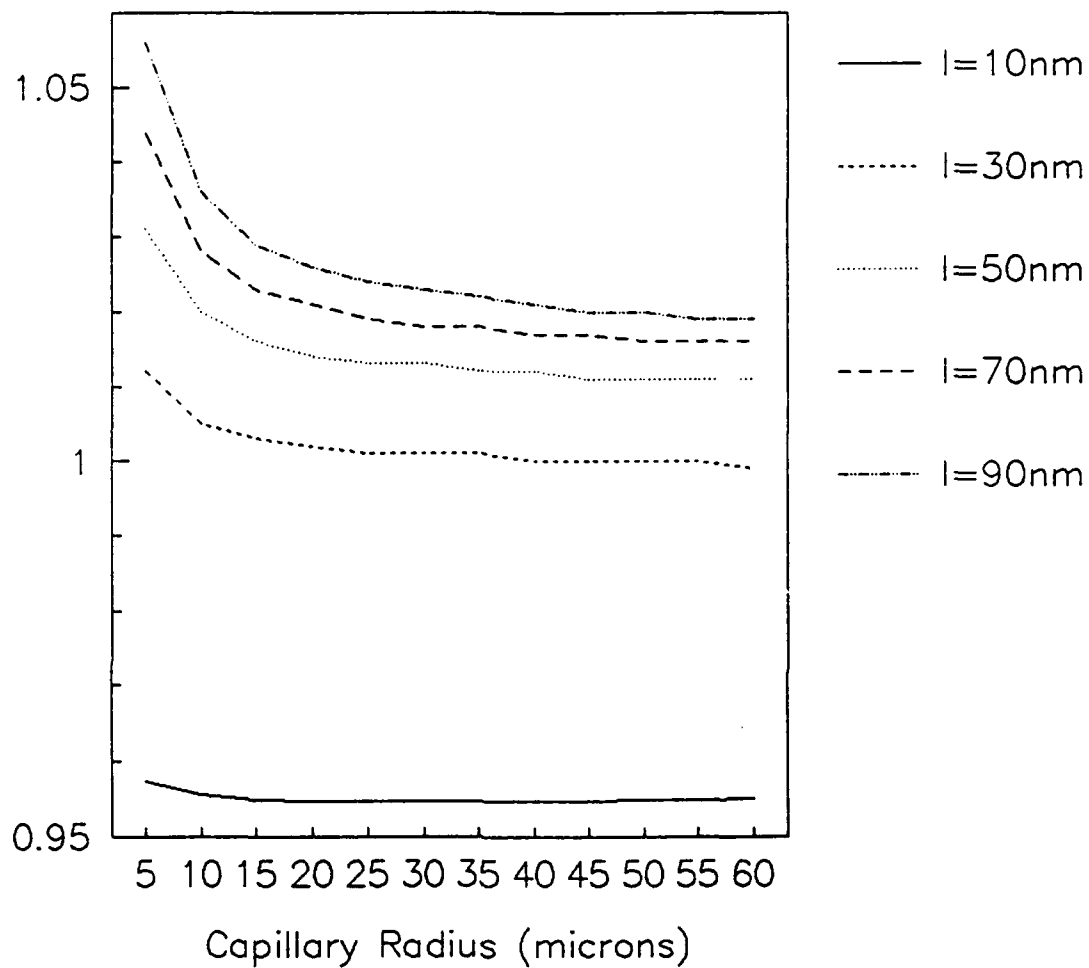
Figure 4.4.   Ratio of percent deviations of the van der Waals interaction from the flat plate results as predicted by the Hamaker theory and surface mode analysis.

1990 RESEARCH INITIATION PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.

FINAL REPORT

MULTI-BODY DYNAMICS EXPERIMENT DESIGN

Prepared by: Roger C. Thompson, Ph.D.
                      Assistant Professor

                      Sandra Scrivener
                      Graduate Research Assistant

Department and   Aerospace Engineering
University: The Pennsylvania State University

Research Location: Astronautics Laboratory/VSSS
                      Edwards AFB
                      Edwards, CA 93523

USAF Researcher: Alok Das, Ph.D.

Date: 5 July 1991

Contract No: F49620-88-C-0053

ABSTRACT

The design of an experimental facility for the investigation of the behavior of multi-body dynamic systems is presented. The objective of the experimental facility is to provide a structure that can be used for experimental testing of various dynamic analysis and control techniques. Results obtained with this facility can then be compared to the analytical projections, providing a means of evaluating the methods and their numerical algorithms. The apparatus is a two-link flexible manipulator with one degree-of-freedom motorized joints and tip masses attached to each link. The motions of the elements are confined to the horizontal plane; the inner arm is allowed to rotate full circle, but the outer arm is limited to a rotation of ± 130° from the fully extended position. Each arm, or link, is highly flexible with first mode vibrations of approximately 0.5 Hz. Multi-body dynamic behavior is exhibited by this structure because the arms act as flexible connectors between the two bodies (the tip masses) attached to the end of each arm. Friction between the structure and the table which supports it is extremely small because the experiment is floating upon a thin (0.004 in.) film of air with the air bearings located under the two tip masses. A large, one-piece granite table with a highly polished, flat, and level surface supports the apparatus and contributes greatly to the nearly frictionless conditions. The detailed design of the flotation system, motor specifications, and mechanical components are presented and compared to the criteria developed in a preliminary proposal for the multi-body dynamics experiment. Analyses of general design options and the advantages of the chosen configuration are also included. All of the mechanical components are presented as complete, detailed designs in the form of engineering drawings. Finally, the capacity for adapting the structure to experiments with various objectives is examined, and recommendations for future applications of this facility are presented.

ACKNOWLEDGEMENTS

I. INTRODUCTION

The motion of complex vehicles composed of multiple modules connected by flexible components (multi-body dynamics) is a topic of intense interest for developers of the next generation of large-scale spacecraft. Because of the complexity of large space structures, mathematical modeling techniques must rely upon equally complex numerical codes to predict the behavior of the structure. Consequently, the accuracy of the individual codes will depend upon the degrees-of-freedom permitted in the model and any simplifying assumptions a ut the motion and controls that will be encountered. The accuracy of the mathematical models constructed by each algorithm is usually estimated through computer generated simulations. Although some experimental facilities are in place, there has not been a consistent, large-scale, ground-based test-bed available for independent laboratory testing of multi-body dynamics and control technology.

Facilities for testing spacecraft dynamics are usually highly specialized. Because the environment of space cannot be duplicated on the Earth's surface, experiments are designed to examine very specific subsystems or simplified motions. Initial test articles typically consisted of simple structures, usually a rigid central body with flexible elements attached. Examples include the Draper/RPL experiment and the JPL Antenna experiment, both of which were supported by the Air Force. Other recent experiments (Refs. 1-3) in flexible manipulators differ in that they are true multi-body structures, but the purpose of these experiments is to control the position of the tip when the motions are relatively slow.

The objective of this project is to design an experimental facility for the USAF that will be available for various studies in multi-body dynamics. Although the immediate requirement for this test-bed is the evaluation of modeling software developed commercially and sponsored through government agencies, the design criteria also established a long-term goal of providing a suitable vehicle for testing all aspects of multi-body behavior. Individual topics to be examined

will include coupled motions, sensor/actuator dynamics, control/structure interactions, structural identification, and optimal control formulations. Consequently, the Multi-Body Dynamics Experiment will be designed to be highly adaptable such that many different experiments with different goals can be executed on the same apparatus. The structure will have "modular" components that can be exchanged easily; different joints, flexible components and masses can be substituted as the need arises.

The Jet Propulsion Laboratory (JPL) was commissioned by the USAF to define the requirements and determine the global operations characteristics of the Multi-Body Dynamics Experiment. In their report (Ref. 4), authors Scheid and Schliesmann proposed a two-link flexible manipulator moving in a horizontal plane to reduce the effect of gravitational loading. To simulate, as close as possible, the free motion in space, the structure will be supported upon a cushion of air over a highly polished, extremely flat granite table (see Fig. 1). The JPL report then proceeds to define the baseline requirements for the two-link structure, the control computer, and the sensors.

A summary of the proposed structural and performance parameters are shown in Table 1. The JPL report included three prospective designs with principal natural frequencies of 0.5 Hz, 0.75 Hz, and 1 Hz, and the 0.5 Hz design was chosen for the Multi-Body Dynamics experiment.

Table 1

Initial Design Parameters

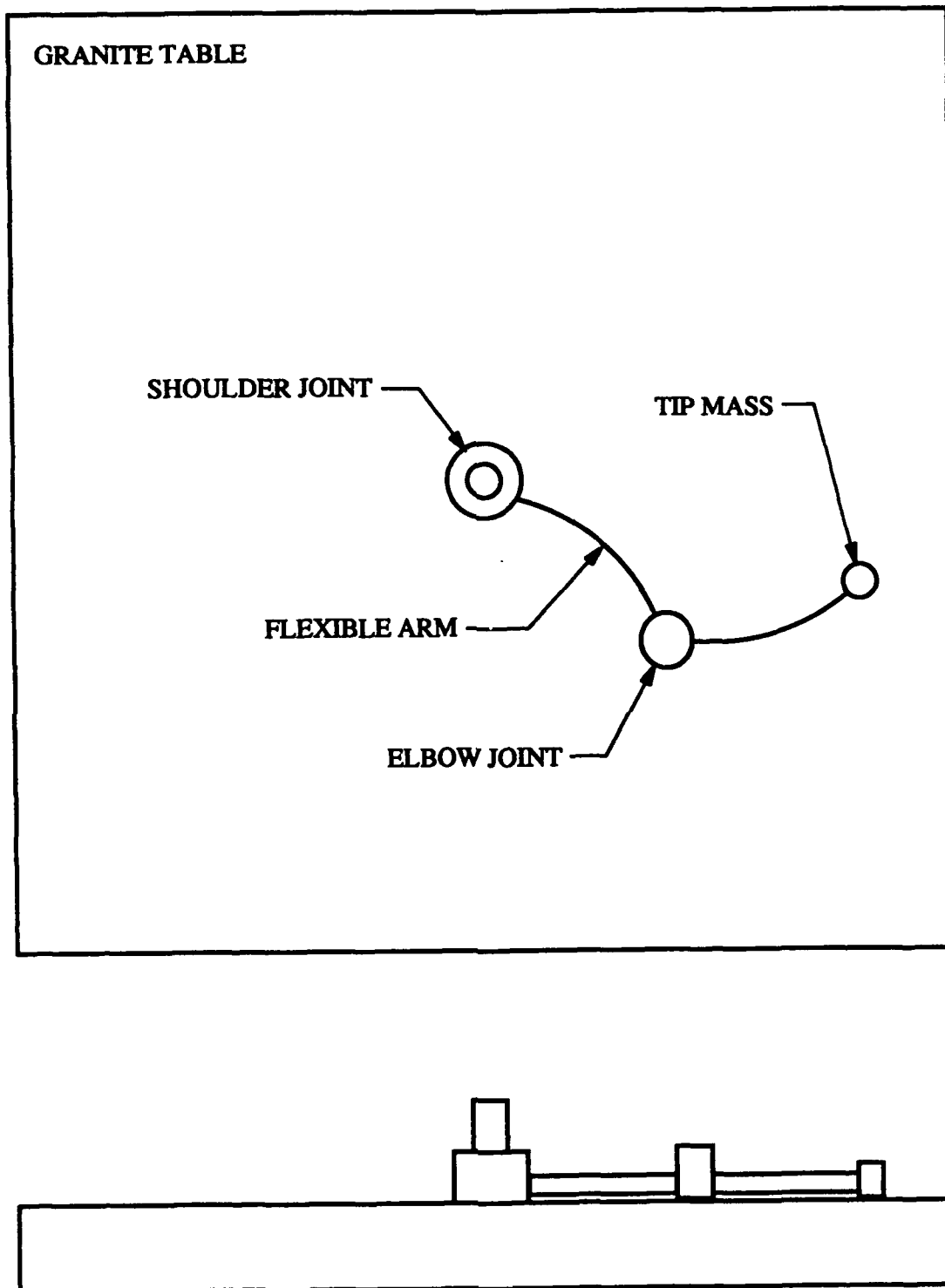| Flexible arm length:	1 m	Beam mass:	0.5 kg
First mode frequency:	0.5 Hz	Shoulder motor torque:	31.4 Nm
Elbow motor torque:	6.2 Nm	Angular velocity:	3 rad/s
Angular displacement:	135°		

Figure 1   Initial Configuration

The most critical components of this structure are the flotation system and motorized hubs; they must be specified in order to complete the design of the mechanical connections, mounting plates, and tip masses. In fact, the flotation system dominated the earliest activities on this project because the lifting capacity of the system would impose constraints on the allowable motor mass and torque. After completing the research on the flotation system, the motors were specified and initial configurations of the hubs were determined. Finally, the remaining mechanical components were completed. In every step of the process, tradeoffs had to be accepted because the initial concept could not be produced with currently available components. Examples of these tradeoffs are included with the detailed systems designs presented in later sections of this report. As each tradeoff was accepted, the impact to the experiment as a whole was examined through analytical simulations.

II.   OBJECTIVES

The objective of this project is to design a suitable experimental test-bed for various studies in multi-body dynamics. The simplest apparatus that exhibits the dynamic behavior common to multi-body structures is a two-link flexible manipulator. It is a structure on which many different experiments (with different goals) can be accomplished. The long-term goal for the project is to test all aspects of multi-body behavior. Individual topics to be examined will include, but not be limited to, coupled motions, sensor/actuator dynamics, control/structure interaction, and time-optimal control technology.

The final product of this project is a complete set of specifications and detailed design of the components that make up the Multi-Body Dynamics Experiment including the fabrication drawings. Because the overall dynamic behavior of the structure is dependent upon the interactions between each element, the baseline structure design must be iteratively refined until complete specification. are achieved. A more typical approach, in which each component is

developed separately would not be as effective as the parallel process that was adopted for this project.

III.   FLOTATION SYSTEM

Experiments were performed during the 1989 USAF-UES Summer Faculty Research Program on a prototype flotation system for the Multi-Body Dynamics Experiment (Ref. 5). A flotation pad was constructed with removable distribution manifolds, made of acrylic (plexiglass), so that different patterns and sizes of vent holes could be tested with low cost components. Lifting capacity, stability under high speed motions, and air supply requirements were measured for the different manifolds. All tests were conducted on a granite table with a surface finish equivalent to th.. specifications of the table to be supplied for the Multi-Body Dynamics facility. Calibrated masses were loaded onto the pad as equally distributed as possible, and the gas pressure was increased until the pad was floating freely.

The results indicated that vent area was not necessarily the factor that determined the lifting capacity of the pad, but it did greatly affect the line pressure required to lift a given mass. A manifold with 36 small holes (0.8 mm diameter) distributed around two concentric rings performed much better than a 3-hole manifold with substantially larger (3 mm diameter) holes even though the vent area of the 36-hole design was 25% less than the area of the 3-hole design. A second 36-hole plate, with 1.2 mm diameter vents, was fabricated and tested against the similar design with smaller vents. The results were quite surprising; the manifold with small vents required 10 psi of line pressure to lift 3 kg whereas the new prototype required only 1.7 psi to lift the same mass.

In every trial with different mass loadings the air gap was measured at 0.1 mm (0.004 in.). Furthermore, the flow rate measured at the compressed air cylinder used for these experiments indicated that 16 L/min was required to lift 3 kg for both 36-hole manifolds. The flow rate is

proportional to the mass lifted, and not the plate design, because the air gap and exit pressure are the same for any free floating mass. The plate design does, however, limit the mass that can be lifted because it limits the volume flow rate that can be achieved.

The performance of the 36-hole manifold design was so successful that it was accepted as the conceptual design of the flotation system, but the mass of the prototype (836 gm) would be too large to incorporate into the final design of the structure. Consequently, the pad needed to be reshaped to reduce the overall mass.

Because the line pressure is so low, the flotation pad does not require thick walls, heavy fasteners, or high strength materials. An all acrylic design was considered, but the potential for warping was too great; most low density plastics also warp quite easily. The prototype flotation pad was an aluminum base plate with an acrylic manifold bolted to the aluminum. A test of an all aluminum manifold revealed a tendency to vibrate, or chatter, on the granite table if the air pressure was not carefully controlled. The acrylic plate could accept wider variations in air pressure before chatter was induced. The apparent vibration damping characteristics of the acrylic are a desirable aspect of this design. In addition, the aluminum base plate provides a stiff surface that will reduce any tendency to warp.

It was decided to retain the aluminum/acrylic combination, but simply reduce the thickness of both components to lower the mass of the system. The acrylic plate can be bonded to the aluminum thus eliminating the mass of the bolts. The final design of the assembled flotation pad is shown in Fig. 2. Note that Fig. 2 shows the flotation pad as viewed from underneath with the acrylic plate in place, and the figure also provides a cross-sectional view to show the stiffeners and bonding surfaces.

Aluminum Backplate

Acrylic Plate

12 Holes
Equally Spaced
on 1.5" Dia.

24 Holes
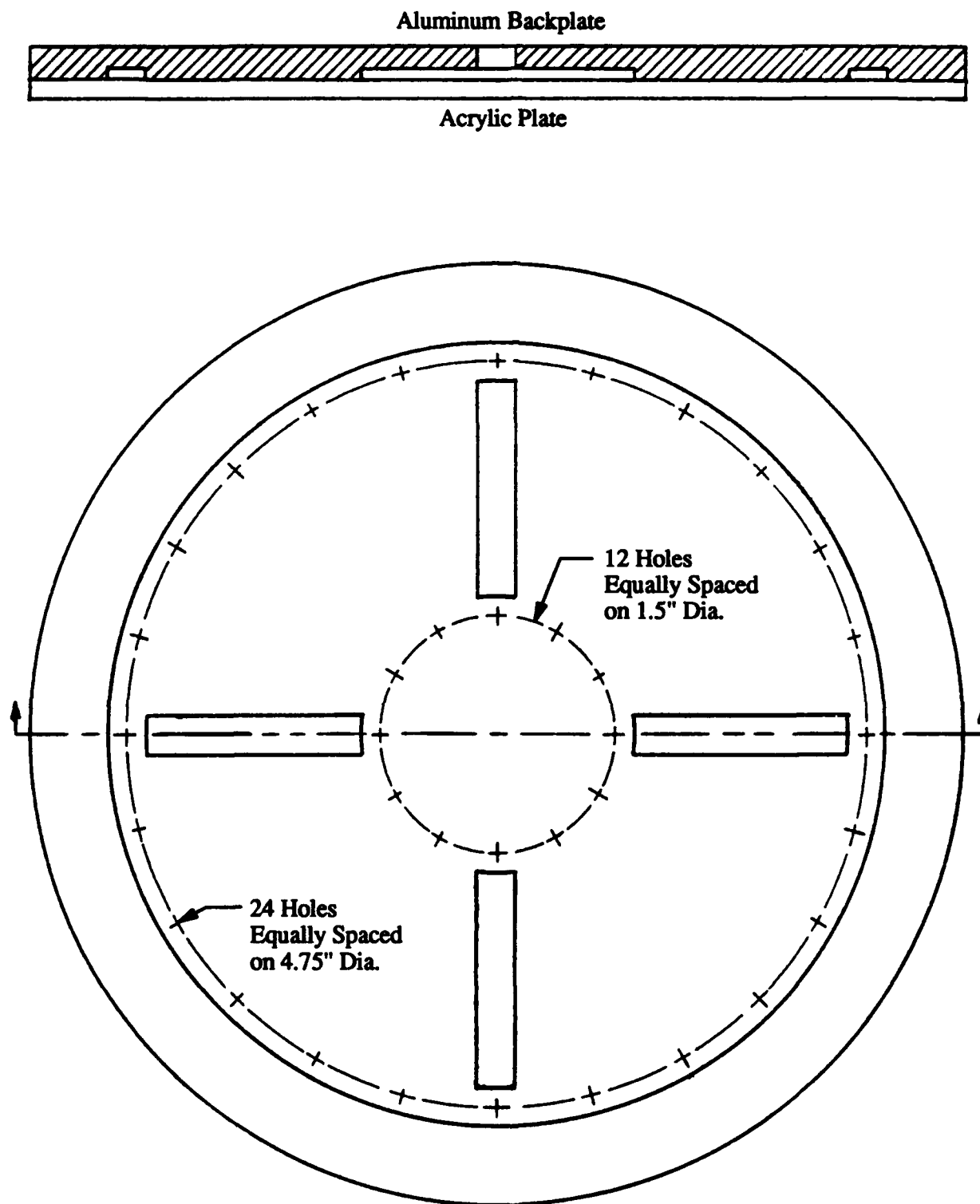Equally Spaced
on 4.75" Dia.

Figure 2   Flotation Pad Final Design

Several options for delivering air to the two flotation pads were considered. On-board compressors are attractive because the effect on the dynamic behavior of the structure is easy to predict; the mass and inertia properties are well known and constant. Commercially available compressors that can deliver 16 L/min at 3 psi invariably exceed 3 kg, and compressors that are less than 3 kg have typical volume flow rates of 2 L/min at up to 5 psi. On-board gas cylinders are the next best choice, but the mass varies as the gas is expelled. A cylinder 10 cm long and 7 cm in diameter with an initial pressure of 1800 psi has a mass 2 kg not including the pressure regulator. Unfortunately, it can supply air to the flotation pad for only 30 sec when lifting a 3 kg load.

Therefore, it will be necessary to use a central compressor or compressed air cylinder and distribute the air to the flotation pads with flexible tubing. The tubing will affect the dynamics of the structure causing frequency shifts and an increase in the damping factor, but experiments have shown (Ref. 5) that the shifts are very small. The effect of the tubing will be measurable, but it will not be sufficient to alter the predominant behavior of the structure.

IV.   MOTORIZED JOINT DESIGN

The most critical components affecting the dynamic behavior of the structure are the motorized joints at the shoulder and elbow. Each joint must rotate smoothly with low friction, the rotary inertia of each joint must be relatively small, and the mass of the elbow motor must be relatively small. As the mass of the elbow motor gets larger, the dynamic behavior of the structure approaches that of a pinned-pinned beam, and as the rotary inertia increases, the behavior approaches that of a fixed-fixed beam. In other words, as the mass and inertia of the joints increase, the apparatus behaves less like a multi-body structure and more like a simple beam structure. Therefore, the mass limits prescribed in the preliminary design study must be maintained if possible.

The combined constraints of low mass, small size, and high torque were difficult to achieve. The original concept for the structure included direct drive, frameless DC motors with hollow shafts; the stator would be the outer portion of the joint and the rotor would be the inner part. It was anticipated that the frameless, hollow-shaft style of motor would have substantially less mass and rotor inertia than is found in their framed, solid-shaft counterparts. After examining several manufacturer's catalogs, it was obvious that a motor with the required characteristics could not be found.

To achieve high torque at low speed, the magnetic field strength of a motor must be increased; more magnets must be used and the magnets must be larger and heavier. Consequently, the mass of the motor increases substantially for a modest increase in torque. A second disadvantage of the higher field strength is that the motor exhibits a jerking motion at low speed known as "cogging." The ferrous rotor tends to align itself with the closest magnet and to resist being displaced from that alignment.

The high mass of the motors alone (up to 20 kg) eliminates them from consideration for this experiment. Because the investigations to be performed with this structure will rarely require more than one revolution from a direct drive motor, cogging torques will be a significant problem; the flexible modes will be disturbed and precise positioning of the beams will not be possible. To reduce the mass and to achieve the desired torque, a gear reduction unit must be included. After an extensive search of manufacturer's literature on high reduction, low mass, and low backlash (freeplay) gear drives, only one acceptable model was found for each hub. Both gearmotors have zero cogging, and the continuous stall torque is equal to the torque specified in the initial design report. Both motors have an optional tachometer and optical encoder, and the elbow motor has, in addition, an electrically released brake. A summary of the pertinent motor performance characteristics are given in Table 2.

Table 2

Motor Performance Ratings

	Shoulder Motor	Elbow Motor
Make and Model	PMI S9M4H/H9D60	PMI S6M4H/H6D60
Rated Speed	50 rpm	50 rpm
Continuous Torque	31 N·m	6 N·m
Peak Torque	64 N·m	34 N·m
Mass	5.93 kg	3.26 kg
Moment of Inertia	1.13×10^{-4} kg·m^2	2.90×10^{-5} kg·m^2

The mass of each motor is, however, higher than the estimates in the original design criteria. It was then necessary to determine how the increased mass would affect the dynamic performance of the structure. A simple analysis, using the NASTRAN structural analysis program, was performed. The results indicate that the natural frequency spread was not widened for the 0.5 Hz design, but the modal frequencies for the higher modes of the 1 Hz design were pushed beyond acceptable limits. This result justifies the choice of the 0.5 Hz design as the preferred configuration of the experiment.

The motors that were selected are the result of a tradeoff. The prescribed torques could not be achieved without a gear reducer, but the gearmotor acts as a rotary damping device when the motor is not powered. Consequently, there will always be some resistance to free rotation of the hubs. In addition, the motors are relatively long with a small diameter. This raises the center of mass of the hubs which can, in turn, subject the structure to torsional loads if the flotation pads are not large enough to stabilize the mass. Finally, the elbow motor is slightly heavier than the original design specifications. Additional experiments found, however, that the flotation system could lift the additional mass without modification (Ref. 6).

V. MECHANICAL COMPONENT DESIGN

Once the air support pads and motors were defined, designing the remainder of the mechanical components was straightforward, although there was significant latitude in the form and dimensions that were acceptable. There were, however, several characteristics that were important to the design. The elbow motor should be mounted such that its center of mass is as close to the table as possible. In addition, the beams that make up the links should also be kept close to the table surface. This will reduce the in-plane torsional moments developed as the beams flex and reduce the tendency to "tilt" the air supports. Because the beams may have some twist along their longitudinal axis, a provision for relieving the torque should be included, and the connectors at the end of each beam must be as stiff as possible to assure a clamped end condition. Of course all of these conditions are constrained by the mass limitations.

A. ELBOW HUB

The elbow joint is supported by one of the two flotation pads with the motor casing attached to the inner arm and the motor shaft attached to the outer arm. The motor could be mounted with the shaft towards the table or with the casing towards the table. The second choice has the advantage of a lower center of mass, but this causes the beams to be very high above the table due to the motor's total length of over 25 cm. In addition, this configuration would be more difficult to supply with electrical power because the motor case would have to be attached to the outer arm. Consequently, the inverted (shaft down) motor mount was chosen.

The design of the hub could then be completed by starting from the flotation pad and working upward. The element to which the outer beam is attached is called the "forearm hub." It has a largely circular base to which the flotation pad is bonded, a collar that accepts the elbow motor shaft, and a surface to which the outer beam is attached. In order to provide relief for any torsional loads induced by the beam, the beam attachment will be through a roller bearing

mounted in the forearm hub; the details of this attachment will be presented in a separate section. A diagram of the forearm hub is shown in Fig. 3 and the detailed engineering drawings are included in Appendix A.

The inner portion of the hub assembly, called the "elbow motor mount," is designed in a similar fashion. The motor is mounted on the upper side of a flat plate and an identical beam attachment (through a roller bearing) is provided at a right angle and at an equal height above the table. The diagram for the elbow motor mount is shown in Fig. 4 and the fully dimensioned drawings are also included in Appendix A.

The dimensions of the components were based primarily upon the dimensions provided by the motor manufacturer. The beams will be centered on the beam attachment points. This will allow the use of beams with various widths because the attachment point is 4.3 cm. above the surface of the table. A schematic of the elbow hub assembly is shown in Fig. 5.

B.   SHOULDER MOTOR HUB

After completing the elbow hub design, the shoulder motor hub was very simple to complete. The height for the beam attachment, established by the elbow hub, must be matched in the shoulder hub. Because the shoulder motor is mounted below the surface of the granite table, the shoulder hub is attached directly to the shaft. The only additional consideration is that the shoulder hub should be as rigid as possible. Please refer to Appendix A for a detailed drawing of the shoulder hub, but a schematic of the shoulder assembly is given in Fig. 6.
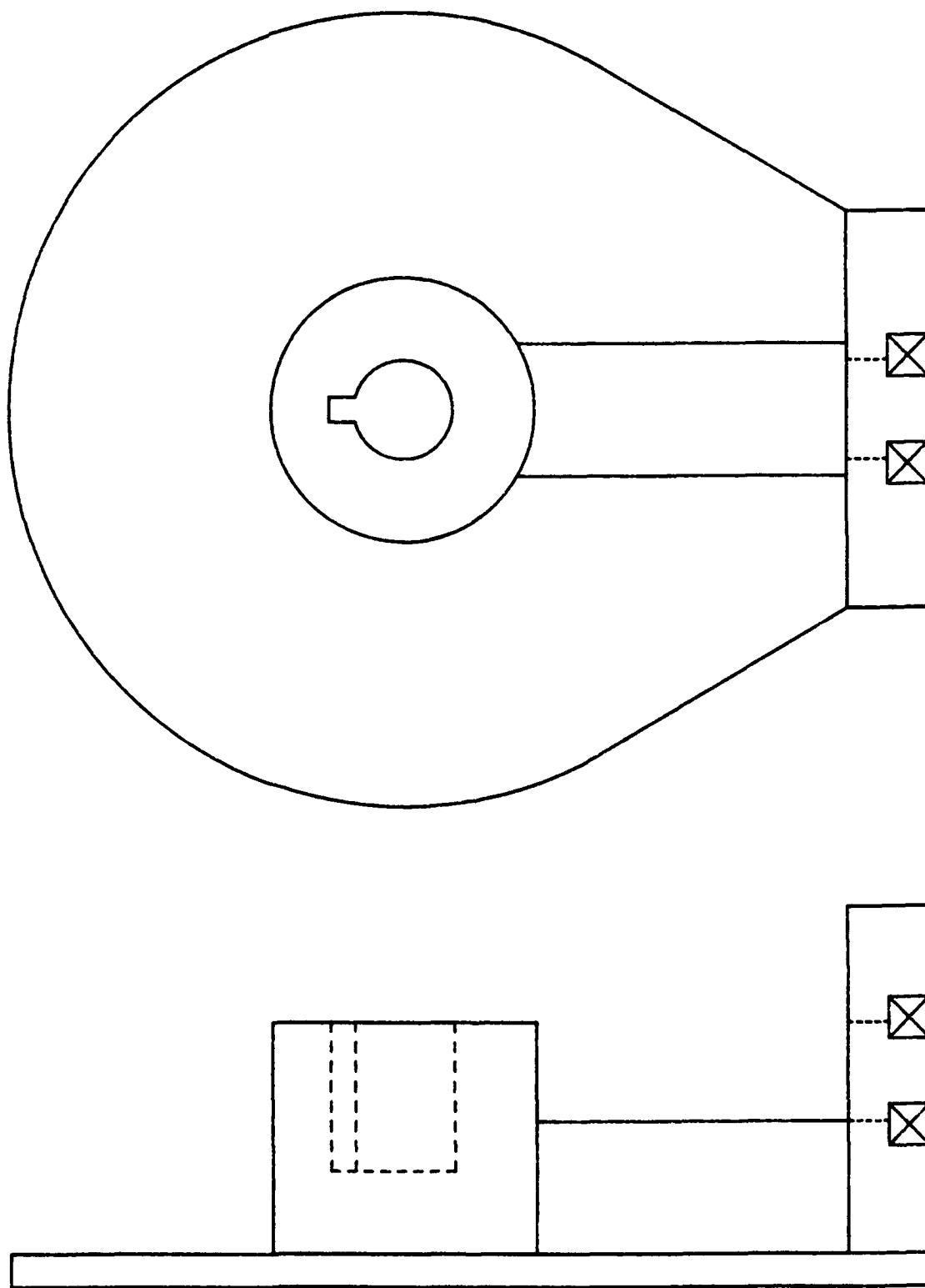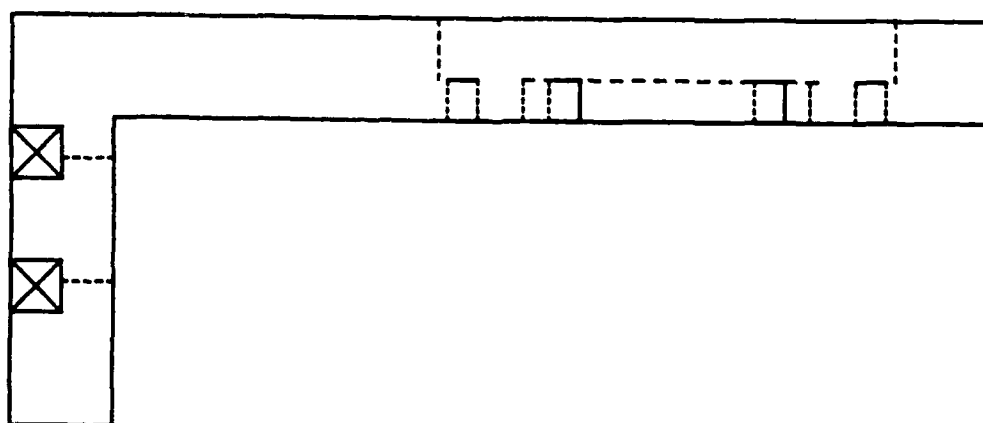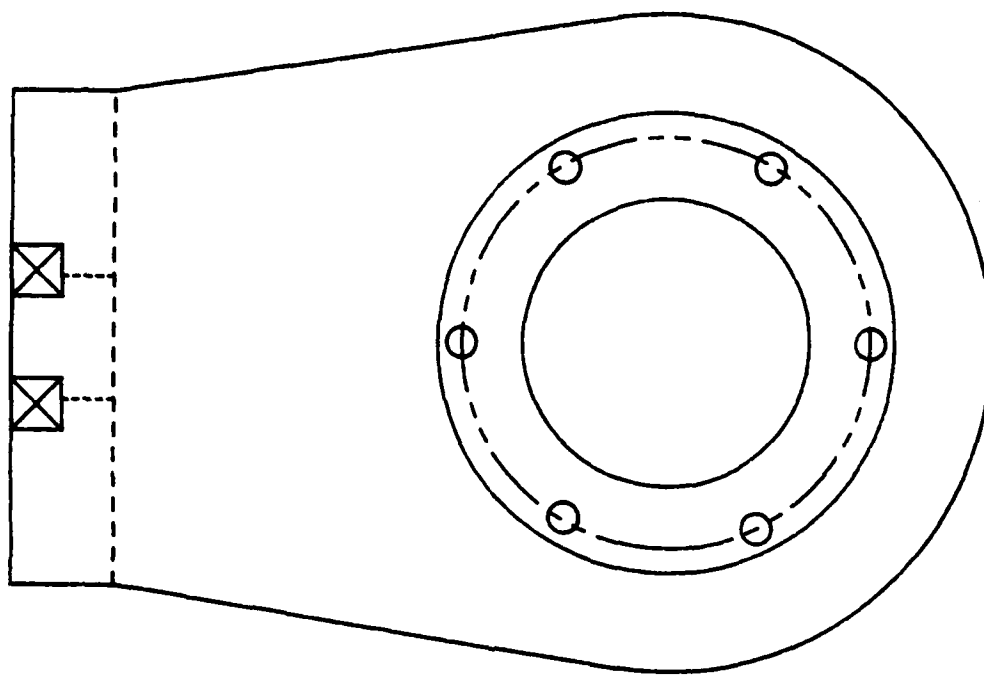
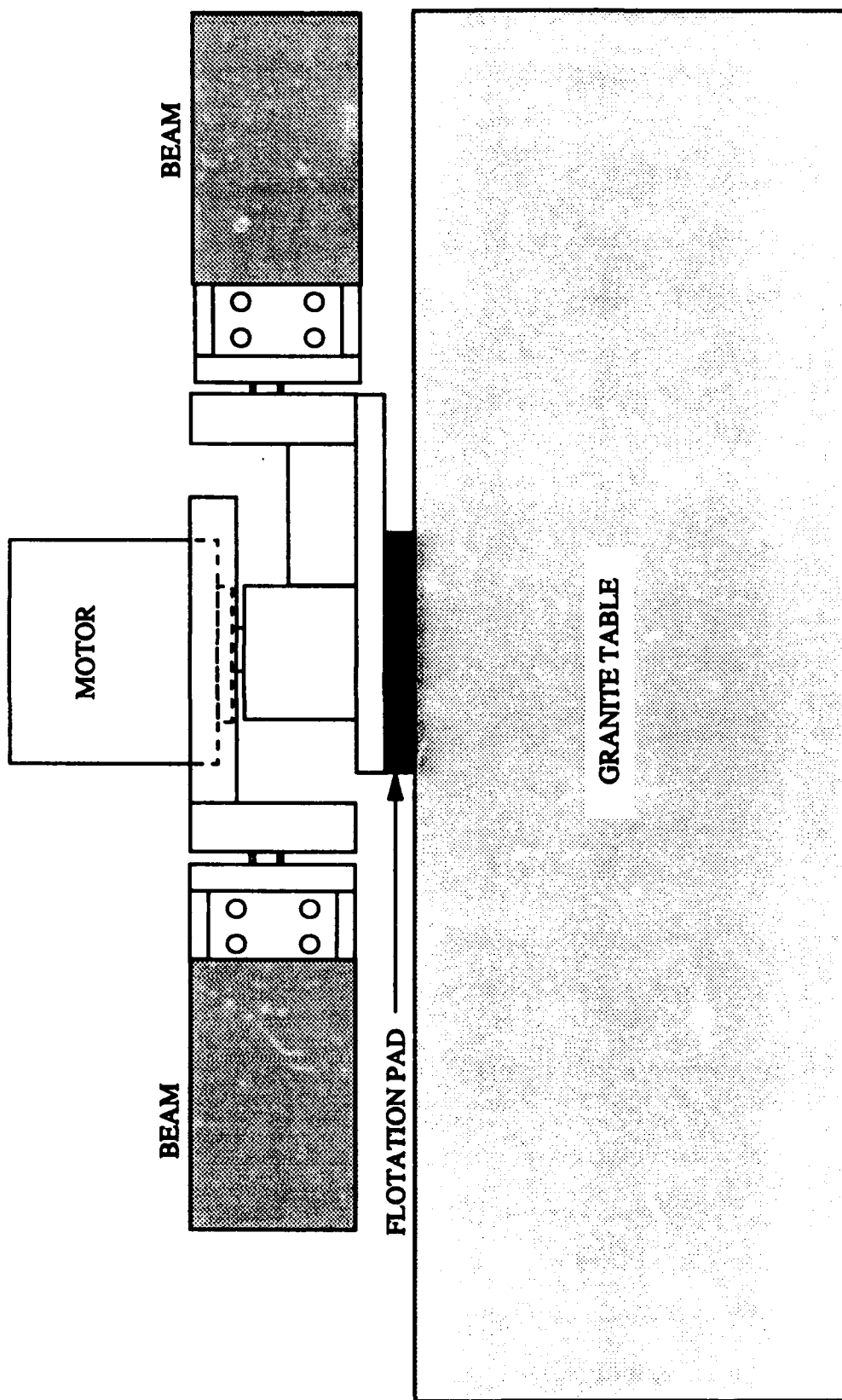Figure 3   Forearm Hub

Figure 4   Elbow Motor Mount

Figure 5   Elbow Hub Assembly
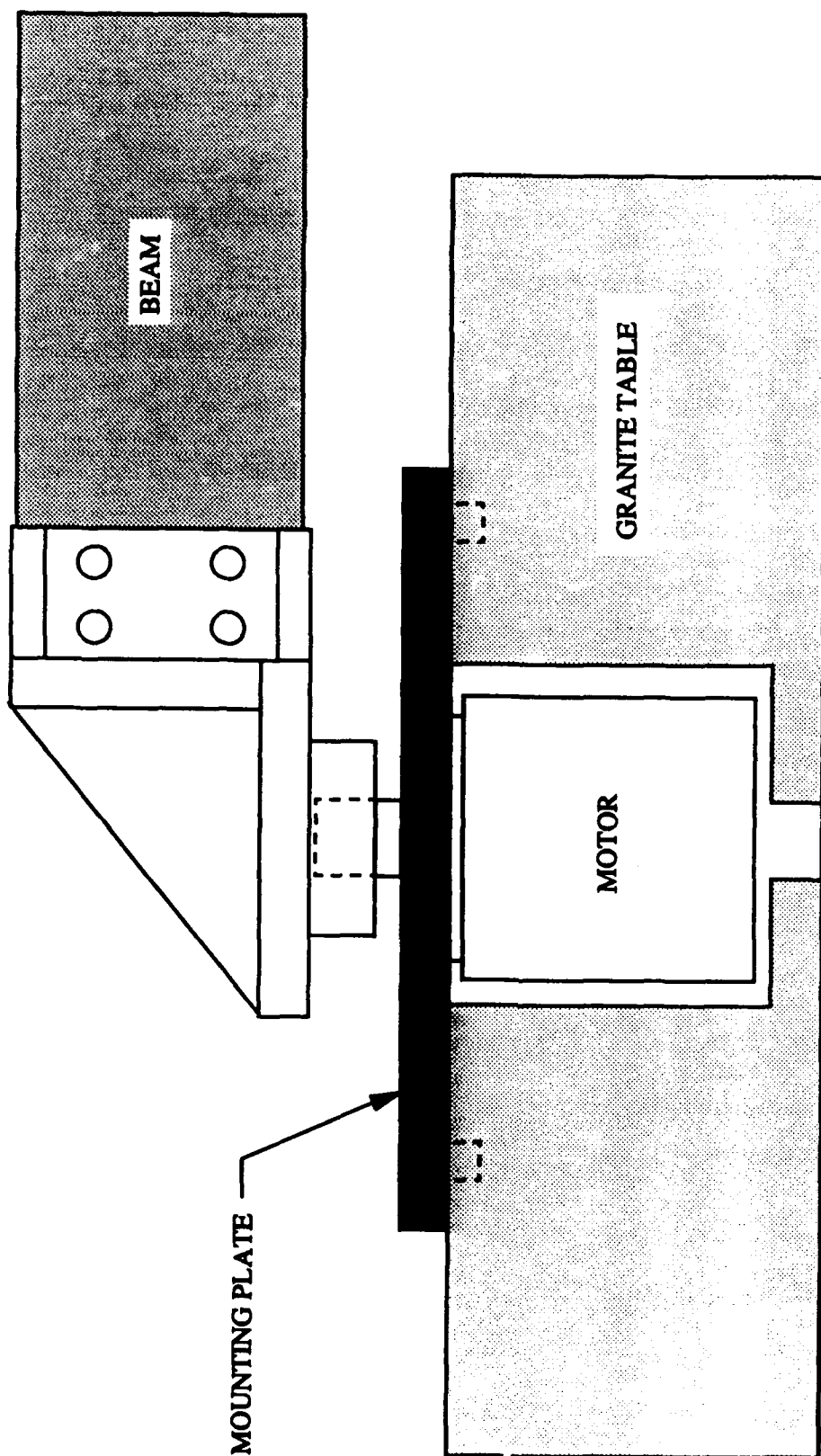
Figure 6   Shoulder Hub Assembly

BEAM

GRANITE TABLE

MOTOR

MOUNTING PLATE

C. TIP MASS

The tip mass is simply a beam attachment, similar to the shoulder hub, mounted above a thin circular plate. A flotation pad is bonded to the underside of the circular plate. Again, the most important criterion in the design of the tip mass is to maintain a stiff structure. In Appendix A, the full scale, detailed drawings of the tip attachment can be found.

D. BEAM CONNECTORS

Each arm of the structure is assumed to flex only in the plane of the table, but in reality, the beams may have a tendency to twist. This torsional moment would tend to tilt the air bearings and cause them to scrape the table top. The air film is only 0.1 mm thick, so the tilt required to force the edge of the air bearing against the table is extremely small. To eliminate this tendency, the elbow hub must be free to rotate about the longitudinal axis of each beam and maintain the cantilevered end conditions in the plane of the table. Therefore, the end of each beam, where it attaches to the elbow hub, will be connected as shown in Fig. 7. The end of the beam is clamped in an end plate with a shaft built into the end. The shaft engages the inner race of the roller bearing with a light press fit and the outer race is press fit into the elbow hub component. A large screw holds the entire assembly together. Although the end of the beam is free to rotate about its longitudinal axis, the shaft prevents rotation of the hub relative to the beam tip in the translational directions. The detailed drawings of the beam end plates are in Appendix A.

SCREW

WASHER

BEARING

BUSHING
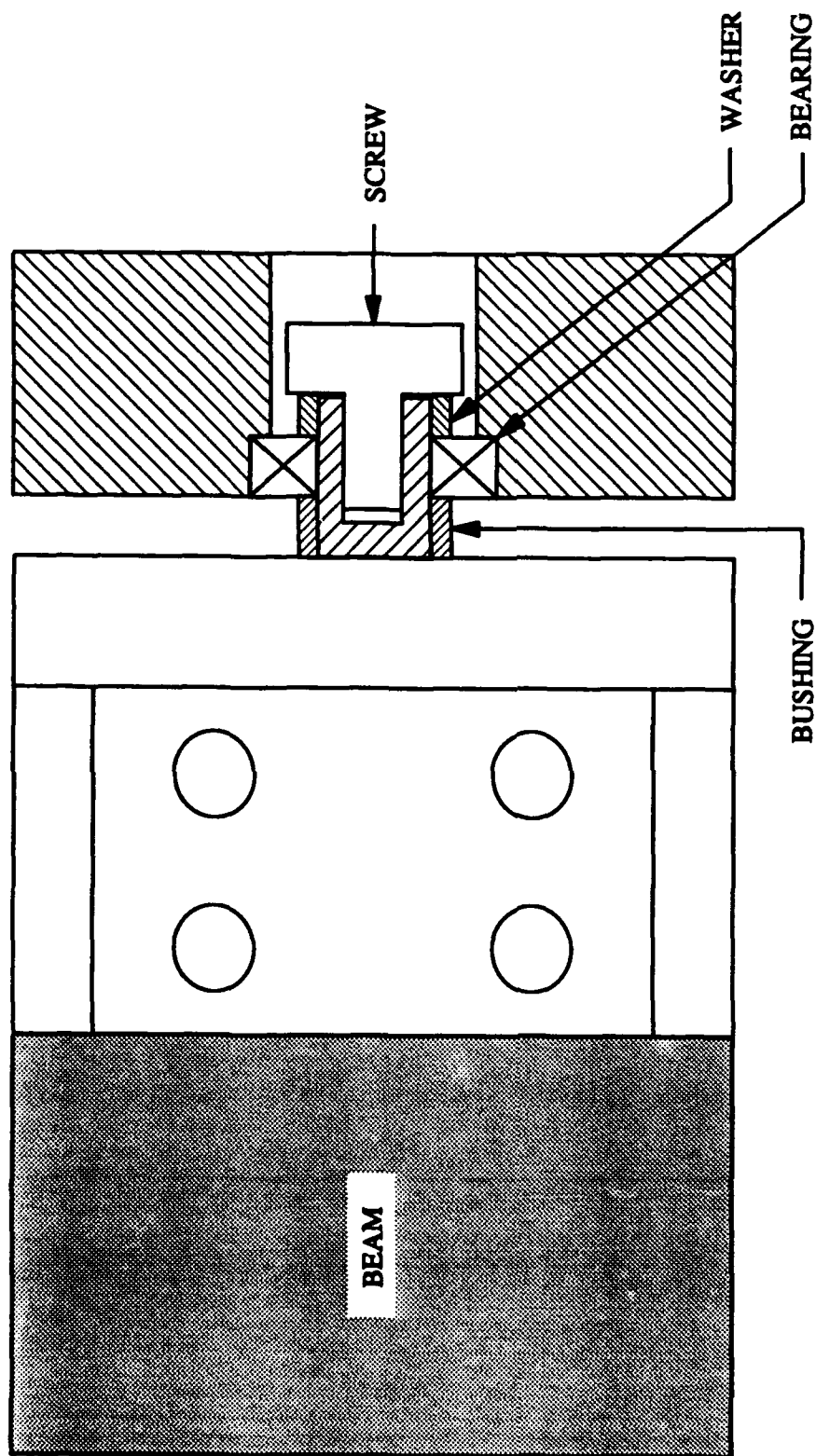
BEAM

Figure 7   Beam Connection Assembly

E.   OVERALL ASSEMBLY

All of the individual components have been designed such that each element can be replaced with similar modules with different dimensions or properties. For example, if an experiment needs an unmotorized elbow hub, only two of the elements, the elbow motor mount and the forearm hub, need to be replaced. Ease of assembly and replacement of components was an important design criterion. All components, with the exception of the beams, will be constructed from 6061-T6 aluminum alloy or its equivalent. The aluminum provides sufficient stiffness, is light in weight, and it is inexpensive. The mass breakdown of the mechanical components are summarized in Table 3.

Table 3

Mass of the Mechanical Components

Component		Mass	Total
Shoulder Assembly	Shoulder Hub	1.0 kg	1.0 kg
Elbow Assembly	Elbow Motor Mount	0.72 kg	5.16 kg
	Forearm Hub	0.44 kg	
	Beam End Plate	2 @ 0.37 kg	
	Elbow Motor	2.34 kg	
Tip Mass Assembly	Tip Mass	0.83 kg	0.83 kg

VI.   BEAM SPECIFICATIONS

The two beams are both approximately one meter in length with a target fundamental natural frequency of 0.5 Hz; the natural frequency for each beam was computed in a cantilevered configuration. The thicknesses specified in the preliminary design (Ref. 4) were dictated by the aluminum sheet metal sizes that are commercially available. The thicknesses which were

considered in the design were from 0.1196 in. to 0.0897 in. Beams with a length of one meter were considered first, with the understanding that the frequency could be raised if necessary by shortening the beam. The tip mass was slightly more difficult to determine. The tip mass for the forearm is the tip assembly, and for the upper arm, it is the entire elbow assembly plus the forearm with its tip mass. Both tip masses, then, depended upon the size of the beams. The beam design process was an iterative one; a size for the beams was chosen, the weight of the mounting assemblies for that size beam was computed, and the natural frequency was calculated with those numbers. Using the results of the frequency calculations, a new size for the beams was chosen and the process was repeated. Rather than go through the design step each time, one initial calculation for the weight of each type of assembly was made to estimate the tip masses. Then the frequencies for both beams for a variety of thicknesses, heights, and tip masses were computed to determine the best configuration.

The analysis quickly showed that most of the configurations had frequencies much higher than the target of 0.5 Hz. There were three mechanisms for decreasing the frequencies (1) increase the beam height, (2) decrease the beam thickness, and (3) increase the tip mass. Ordinarily, increasing the beam length would be an appropriate means of lowering the frequency, but this was not possible because of the size limitation of the granite table. The beam height/thickness ratio was increased until it was felt that a further increase would cause problems with out-of-plane buckling. The tip masses due to the motor/beam mount apparatus on the end of each beam were then recalculated, using the modified specifications, and it was found that these masses were significantly lower than the mass required to produce the target frequency. Therefore, additional weights will have to be added to the tip masses if the frequencies are to be lowered to the 0.5 Hz range.

Once the design criteria with respect to the frequency had been met, it was necessary to consider the strength of the beam and its behavior under loading. The first type of loading

49-23

considered was an axial load due to a constant-speed rotation with the manipulator in the fully extended position. Both the upper arm and forearm were well below the yield stress even when the apparatus is moving at its maximum speed.

The second type of loading investigated was bending stress. A worst case scenario was taken for the analysis: that of a bang-bang maneuver. The shoulder motor has a peak torque of 64 N·m, and the resulting stress on the beam was more than 2.5 times the yield stress of aluminum. A review of the JPL preliminary design indicated that the beam proposed in that report would also yield under the applied load.

Because the bending stress was excessive, and because the apparatus design seemed to fulfill the initial design criteria, it was decided to find another material for the beam rather than drastically change the dimensions of the beam. Note that strengthening the beam by increasing the cross-sectional area would drive the natural frequencies higher. It was difficult to find materials that were strong enough but were not so stiff as to raise the frequency to unacceptably high values. Three candidate materials were found to be strong enough to survive the bang-bang maneuver without increasing the natural frequency too much. Unfortunately, one characteristic all three materials have in common is that they are relatively expensive compared to aluminum.

A.   STAINLESS STEEL

In the first case, stainless steel was substituted for the aluminum beam, but the physical dimensions remained the same. The relationship between the tip mass and the fundamental natural frequency was determined numerically; the result is shown in Fig. 8. The tip masses were then chosen from this data such that the frequencies are near the 0.5 Hz range. The elbow hub would need a total mass of 5.67 kg and the forearm tip mass would need to be 1.36 kg. The material properties, beam dimensions, and natural frequencies are also included in Fig. 8.

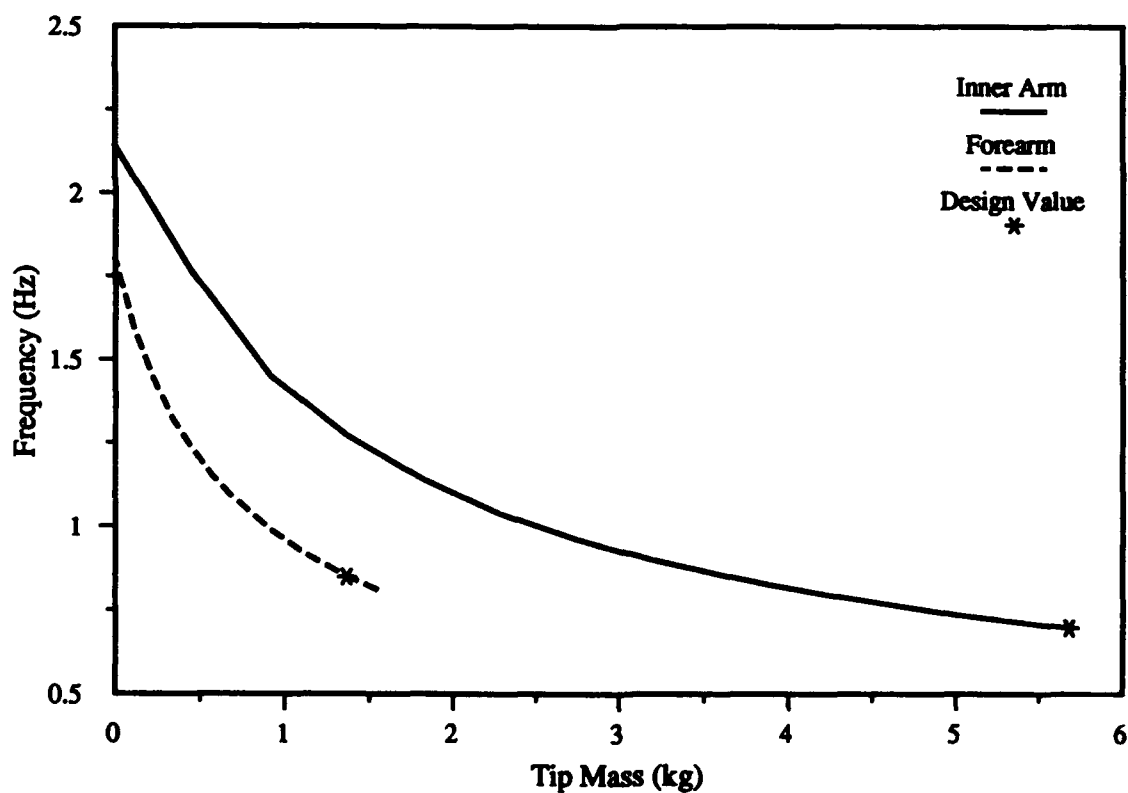| Design Parameters	Inner Arm	Forearm
Length	1 m	1 m
Width	88.9 mm (3.5 in.)	88.9 mm (3.5 in.)
Thickness	3.05 mm (0.1196 in.)	2.28 mm (0.897 in.)
Tip Mass	5.67 kg	1.36 kg
Natural Frequency	0.70 Hz	0.85 Hz
Elastic Modulus	190 GPa (28 x 10^6 psi)	
Yield Strength	520 MPa (75,000 psi)	

Figure 8   Stainless Steel Beam Design

B. TEMPERED ALLOY STEEL

The second alternative material examined was a tempered, quenched alloy steel. The fundamental natural frequency, as a function of tip mass, and the selected structural parameters are shown in Fig. 9. This design permits full operation in the entire range of the motor torques, including peak torque loads, but the mass loading at both the elbow and forearm tip are also large. Therefore, the structure will vibrate in what is essentially a pinned-cantilevered mode.

C. S-2 FIBERGLASS

The third material to be investigated is an S-2 fiberglass composite. This material is readily available in a number of sheet sizes with various thicknesses and layups. It is easily machinable and has the added advantage of being relatively non-conducting. The uni-directional layup was chosen for ease of modeling the mechanical behavior, but another lay-up could prove to be better at satisfying different experimental requirements. When determining the relationship between the tip mass and frequency, it was found that the frequencies were very low even with the lightest mass. Consequently, the beams were shortened to 0.8 m to obtain the proper frequency range. The structural parameters and frequency plot are shown in Fig. 10. This material is capable of withstanding maximum loads from the motors, and the S-2 fiberglass beams are recommended for the Multi-Body Dynamics Experiment.

Design Parameters	Inner Arm	Forearm
Length	1 m	1 m
Width	50.8 mm (2.0 in.)	50.8 mm (2.0 in.)
Thickness	3.05 mm (0.1196 in.)	2.28 mm (0.897 in.)
Tip Mass	5.67 kg	1.36 kg
Natural Frequency	0.55 Hz	0.69 Hz
Elastic Modulus	200 GPa (29 x 10^6 psi)	
Yield Strength	825 MPa (100,000 psi)	
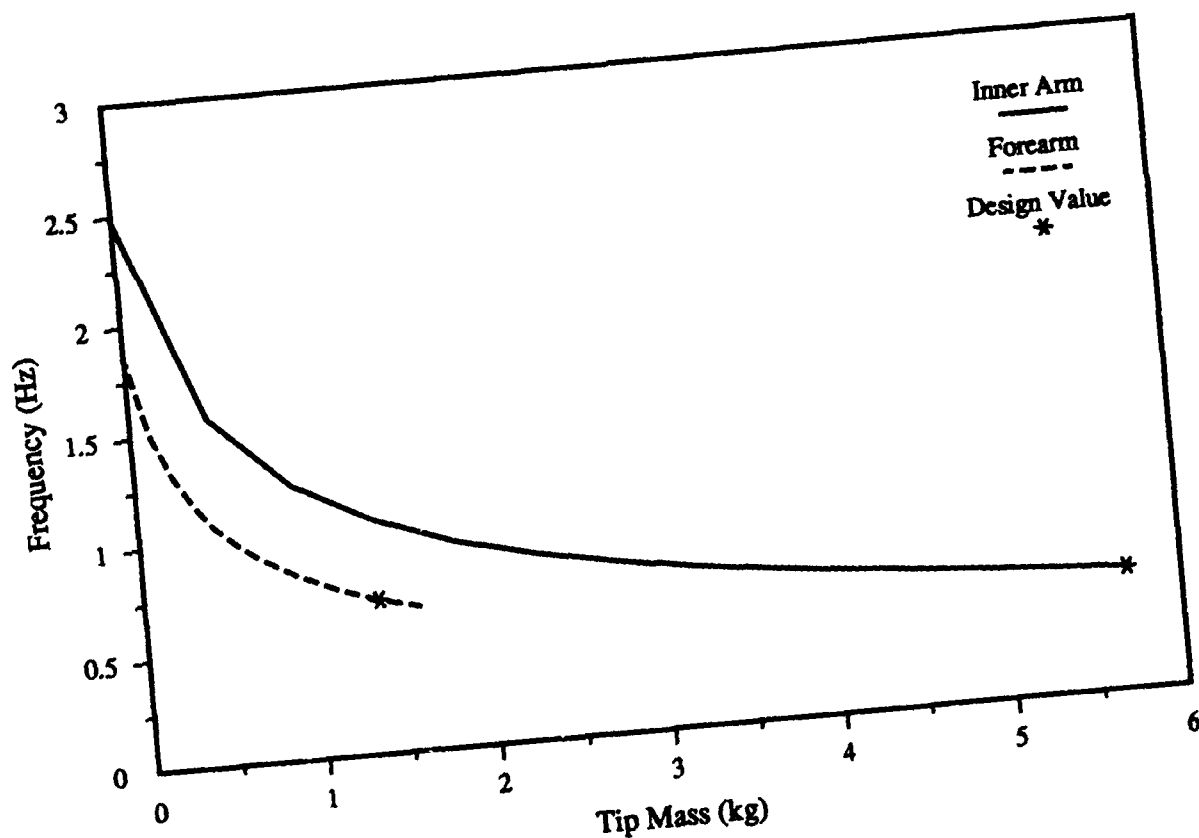
Figure 9   Tempered Steel Beam Design

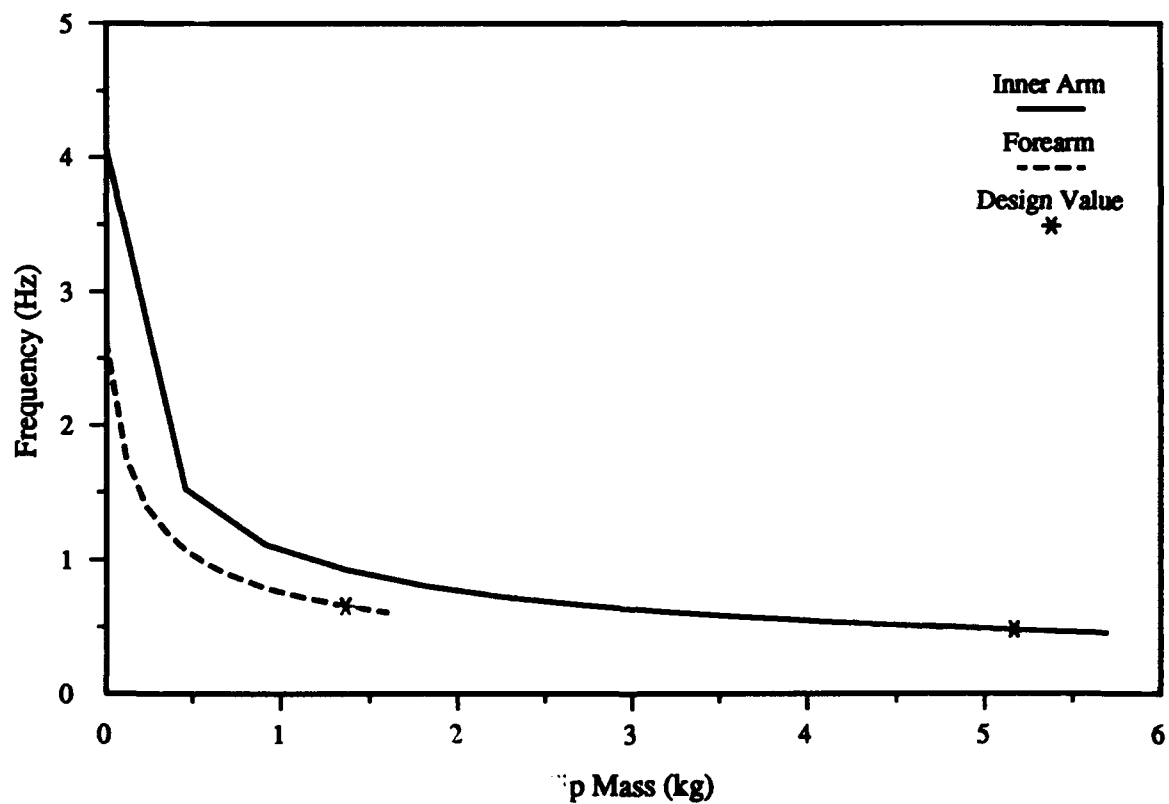| Design Parameters	Inner Arm	Forearm
Length	0.8 m	0.8 m
Width	63.5 mm (2.5 in.)	63.5 mm (2.5 in.)
Thickness	3.05 mm (0.1196 in.)	3.05 mm (0.897 in.)
Tip Mass	5.16 kg	1.36 kg
Natural Frequency	0.48 Hz	0.66 Hz
Elastic Modulus	55 GPa (7.98 x 10^6 psi)	
Yield Strength	1575 MPa (2.28 x 10^5 psi)	

Figure 10   S-2 Fiberglass Beam Design

VII. POWER AND AIR DISTRIBUTION

Air, power, and sensor wire distribution to the structure will be dropped from above the structure to the shoulder hub. From that point, very thin wiring and highly flexible tubing will run along the upper surface of the inner arm to the elbow hub. Sensor leads and the air tubing will be looped across the elbow hub and continue along the upper surface of the forearm to the tip mass. It is obvious that this arrangement will influence the dynamic behavior of the structure, but tests have confirmed (Ref. 5) that this effect is small and it can be easily measured. Because the elbow motor does not rotate more than $\pm 130°$, the loop of wires and tubing across the moving components will not have any significant effect. The wires and tubes dropped to the shoulder hub, however, present a more difficult problem. The suspended cables must be capable of twisting, with minimal resistance, over approximately five revolutions. Fortunately, the length of these cables is sufficient to allow the rotations.

Nevertheless, a much better option is to drop the wires and tubes to electrical and pneumatic rotary couplings mounted on the shoulder hub. These products were examined for potential use in this facility, but the high cost proved prohibitive. An electrical rotary coupling capable of handling the number of leads and current capacity required for this experimental structure would cost over $25,000. Letting the cables twist is not the preferred choice, but the alternative would exceed the budget for this project.

VIII. CONCLUSIONS AND RECOMMENDATIONS

The Multi-Body Dynamics Experiment design has been completed and the finished drawings have been delivered to the Astronautics Laboratory where the facility will be located. Virtually all of the original design specifications were met, although some very specialized materials and products were required to accomplish the task. The design is highly modular such that components can be replaced with elements having different properties and dynamic

characteristics. For example, this structure can easily accommodate imbedded fiber-optic elements (smart structures) for testing and evaluation in structural identification and real-time control applications. The next phase of the project for the Astronautics Laboratory is the fabrication of the components and the assembly of the complete facility.

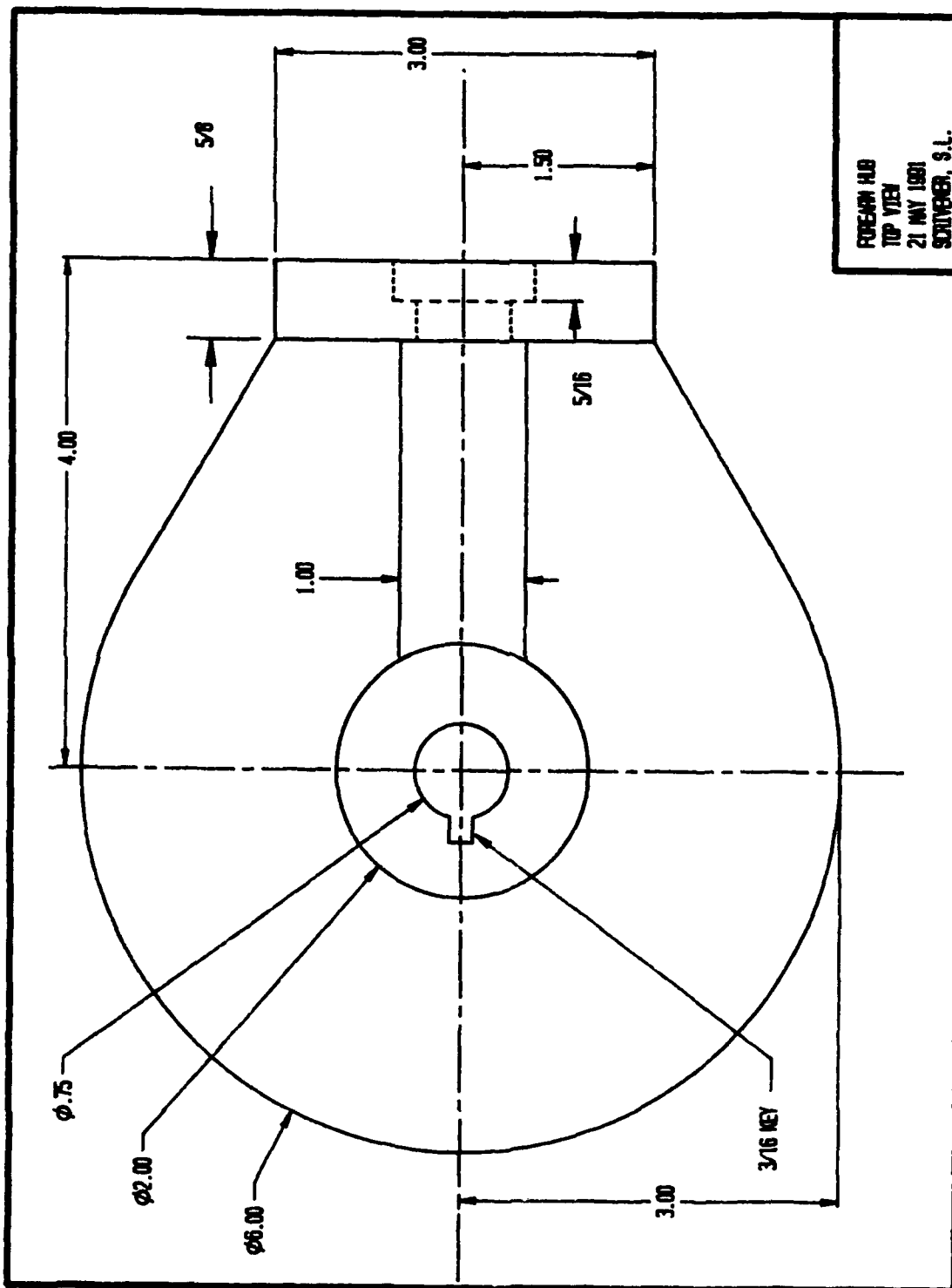The principal recommendation to be made is that a high priority should be given to acquiring electrical and pneumatic rotary couplings. These products will greatly enhance the freedom of motion so desirable in a facility of this type. Secondly, replacing the typical sensor and power wiring with very low-stiffness flat cables, if they become available, would reduce the frequency and damping factor shifts induced by the cables.

REFERENCES

1. Oakley, C. M. and Cannon, R. H., "Initial Experiments on the Control of a Two-Link Manipulator with a Very Flexible Forearm," Proceedings of the 1988 IEEE Robotics and Automation Conference, pp. 996-1002.

2. Buchan, K. S., Carusone, J., and D'Eleuterio, G. M. T., "RADIUS - A Laboratory Facility for the Study of the Dynamics and Control of Elastic Manipulators," Presented at the Seventh VPI&SU/AIAA Symposium on Dynamics and Control of Large Structures, Blacksburg, VA, May 8-10, 1989.

3. Schmitz, E. and Ramey, M., "A Simplified Dynamic Model for an Experimental Planar Manipulator with an Elastic Forearm," Presented at the American Control Conference, Pittsburgh, PA, June 21-23, 1989.

4. Scheid, R. E. and Schliesmann, R. G., "Multibody Dynamics Experiment: Requirements Definition and Preliminary Design," JPL Report No. JPL D-6472, May 30, 1989.

5. Thompson, R. C., "Design and Development of a Flexible Multi-Body Dynamics Experiment," Final Report to Universal Energy Systems, Contract No. F49620-88-C-0053, September, 1989.

6. Thompson, R. C., "Component Design for the Multi-Body Dynamics Experiment," Final Report to Universal Energy Systems, Contract No. F49620-88-C-0053, October, 1990.

APPENDIX A

ENGINEERING DRAWINGS OF MECHANICAL COMPONENTS

FOREARM HUB
TOP VIEW
21 MAY 1991
SCRIVENER, S.L.

3.00

1.50

5/8

5/16

4.00

1.00

Ø.75

Ø2.00

Ø6.00

3.00

3/16 KEY

49-33

FOREARM HUB
SIDE VIEW
21 MAY 1991
SCRIVENER, S.L.

49-34

FIREARM HUB
FRONT VIEW
21 MAY 1991
SCRIVENER, S.L.

$\phi 1.125^{+.000}_{-.001}$

$\phi .75$

3.00

6.00

49-35

DRILL AND C'BORE
FOR 10-32 SOC. HD.
CAP SCREW
6 HOLES EQUALLY
SPACED ON A
2.5 DIA CIRCLE

Ø4.00

Ø1.75

Ø2.00

2.00

4.00

5/8

1.50

3.00

ELBOW MOTOR MOUNT
TOP VIEW
20 MAY 1991
SCHREIBER, S.L.

ELBOW MOTOR MOUNT
SIDE VIEW
20 MAY 1991
SCHWEBER, S.L.

49-37

ELBOW MOTOR MOUNT
FRONT VIEW
21 MAY 1991
SCHOENBERGER, S.L.
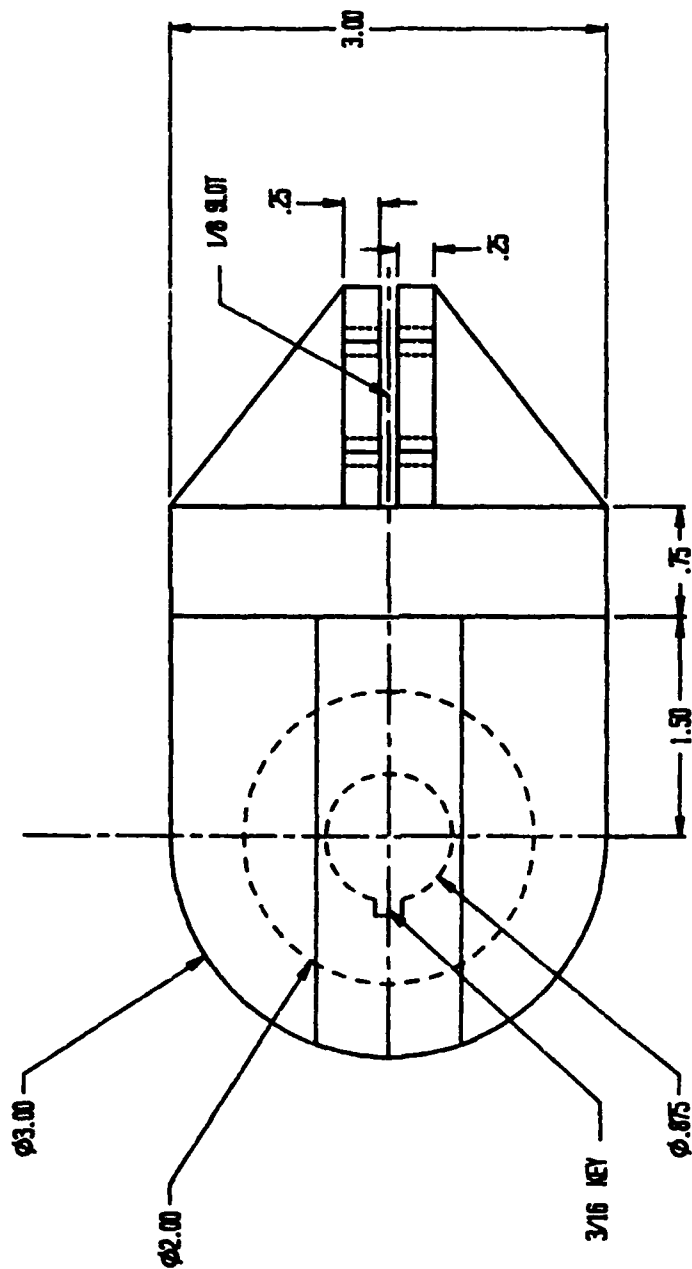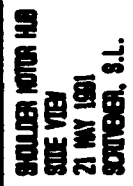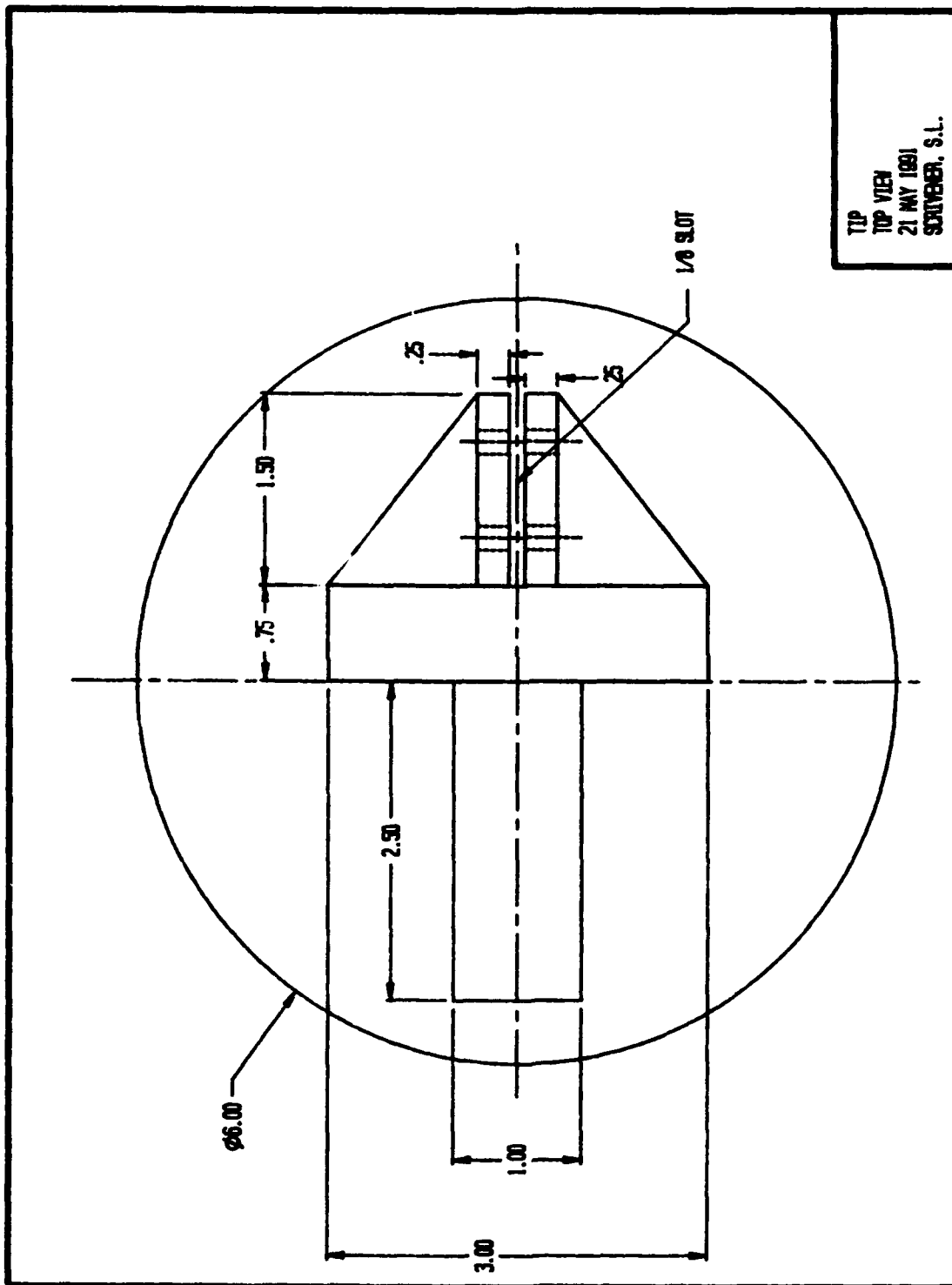
Ø1.125 +.000 -.001

Ø.75

4.00

3.00

1.50

2.50

5/8

5/8

SMALLER ROTOR HUB
TOP VIEW
21 MAY 1991
SCHREIBER, S.L.

3.00

1/8 SLOT

.25

.25

.75

1.50

Ø3.00

Ø2.00

3/16 KEY

Ø.875

SHOULDER MOTOR HUB
SIDE VIEW
21 MAY 1991
SCHRENKER, S.L.

49-40

TIP
TOP VIEW
21 MAY 1991
SCRIVENER, S.L.

Ø6.00

3.00

2.50

1.00

.75

1.50

.25

.25

1/8 SLOT

TIP
SIDE VIEW
21 MAY 1991
SCRIVENER, S.L.

1/8 SLOT

.25

.25

1.50

DRILL AND TAP FOR
5/16 - 18 UNC THD
3/4 DEEP

1.00

3/4

1.50

3.00

BEAM END PLATES
SIDE VIEW
20 MAY 1981
SCHWEBER, S.L.

13/64 DIA

3/8

3/8

.50

.50

.25

.25

1.25

2.50

BEAM END PLATES
FRONT VIEW
21 MAY 1991
SCRIVENER, S.L.

2.50

Ø.50

1.25

3.00

.25

1.50

.25

5/16 DIA
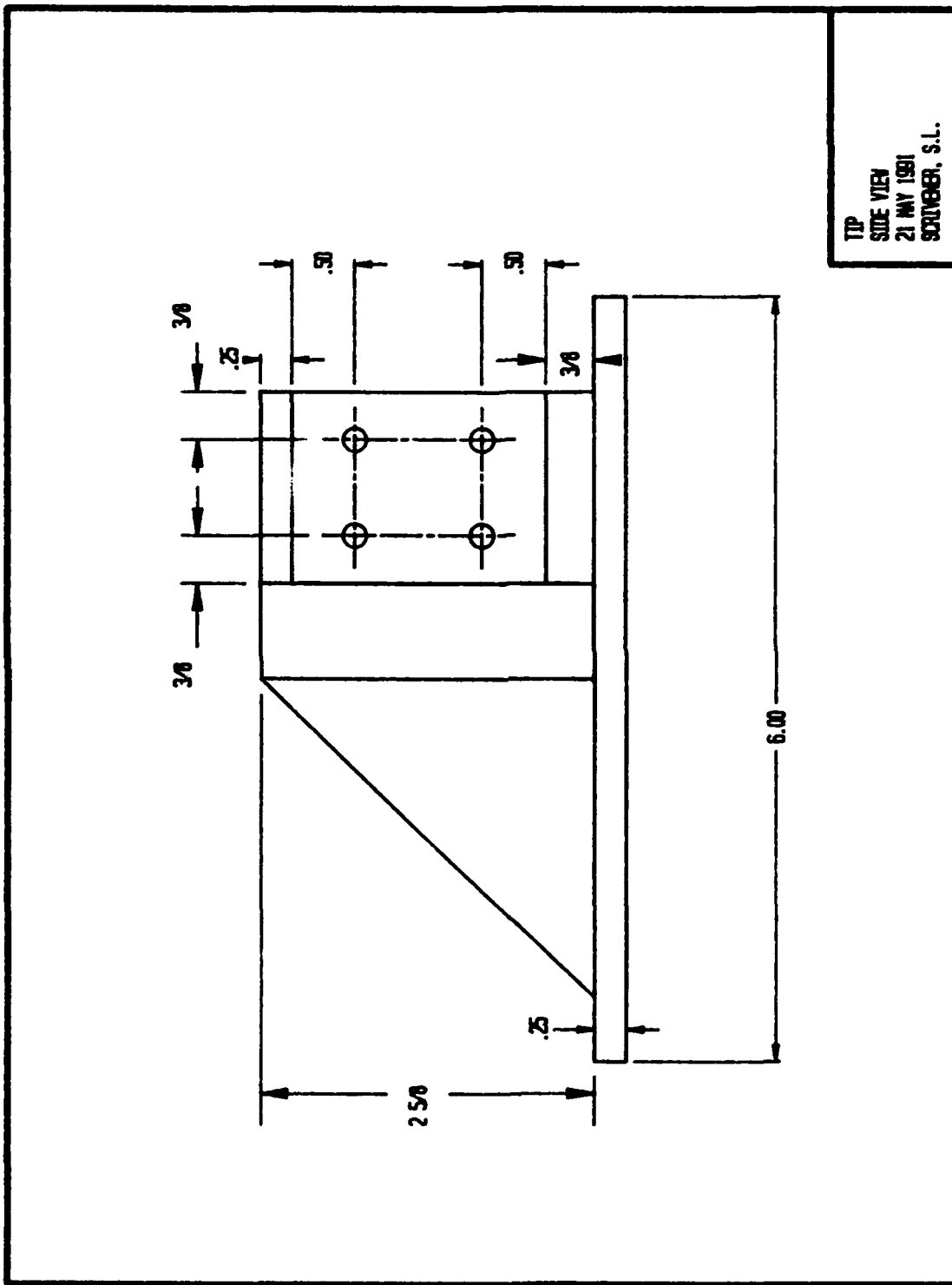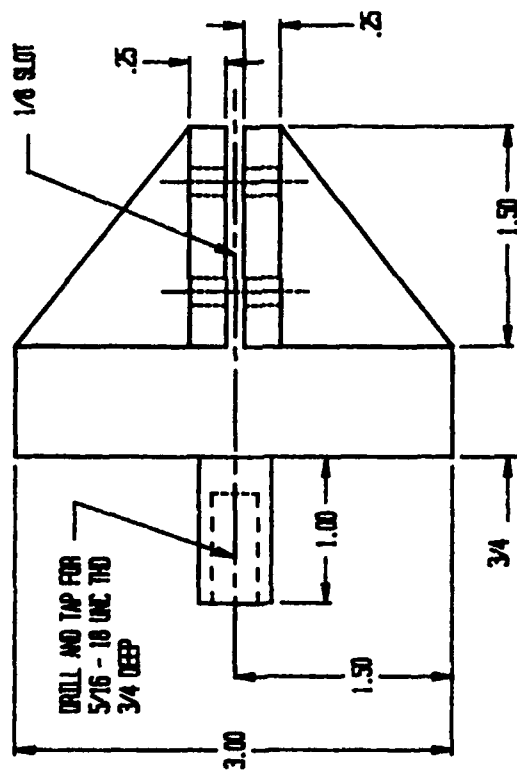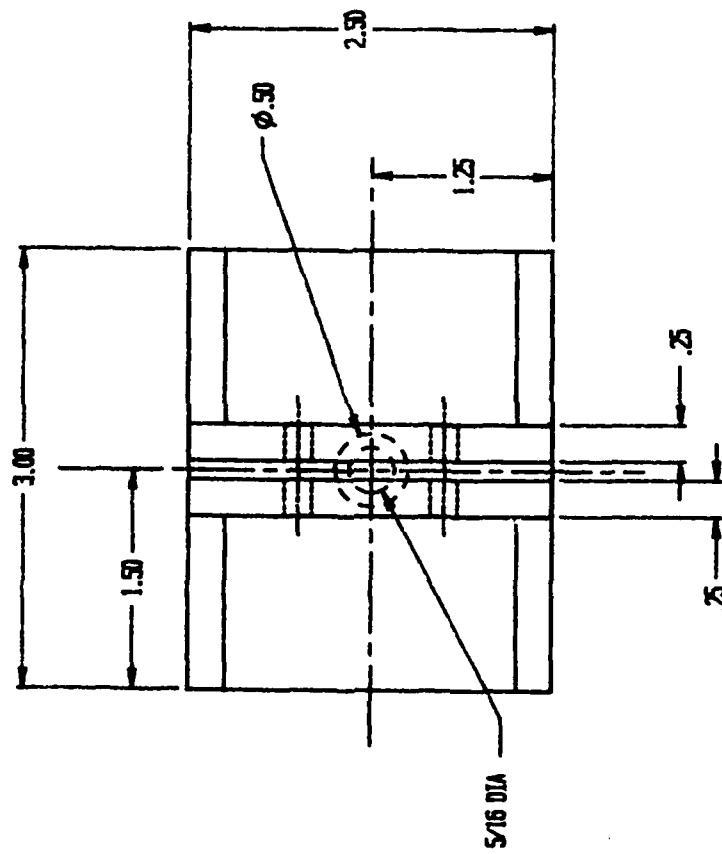
Report # 50
210-9MG-119
Prof. Lane Clark
No Report Submitted

Report # 51
760-6MG-054
Prof. Fabian Hadipriono
No Report Submitted

1989 USAF-UES RESEARCH INITIATION PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by

Universal Energy Systems, Inc.

FINAL REPORT

IMPROVED MODELING OF THE RESPONSE OF PRESSURIZED COMPOSITE CYLINDERS TO LASER DAMAGE

Prepared by:	Harry A. Hogan, Ph.D. Stuart J. Harbert, M.S.
Academic Ranks:	Assistant Professor Research Assistant
Department and	Department of Mechanical Engineering
University:	Texas A & M University
Affiliated Laboratory:	WL/TALE Kirtland AFB, NM 87117
USAF Researchers:	Mr. Jorge E. Beraun Lt. Daniel M. Hodgkiss
Date:	31 December 1990
Contract No.:	F49620-88-C-0053/SB5881-0378

IMPROVED MODELING OF THE RESPONSE OF PRESSURIZED COMPOSITE CYLINDERS TO LASER DAMAGE

by
Harry A. Hogan
and
Stuart J. Harbert

ABSTRACT

The response of pressurized composite cylinders to laser damage is a problem of interest to the Air Force and the Weapons Lab because of its application to the Strategic Defense Initiative. Current numerical models for this problem, however, are generally inadequate in predicting results outside the range of parameters for which a large experimental database already exists. Thus, the advantages of modeling cannot be fully exploited. Several areas for developing an improved model have been identified. The basic goal is to create a more physically based model that includes more detailed deformation and failure mechanisms. This initial effort has focused on a simplified model, the notched tensile bar, which is analogous to the cylindrical pressure vessel and can be studied in extensive detail. Delamination cracks have been treated explicitly using two-dimensional continuum finite elements and the virtual crack closure technique. A series of solutions have been generated for a wide variety of crack locations and lengths. Strain energy release rates are used to characterize the driving force for crack extension for each case modeled. The results clearly demonstrate the value of studying a simplified model to understand the basic trends in the mechanics of problems of this type. A three-dimensional model of the cylindrical pressure vessel will require using lower-order approaches such as shell type elements and the findings of this work will help establish priorities and limits in the range of system parameters needed for such models.

ACKNOWLEDGEMENTS

Sincere thanks to Universal Energy Systems, Inc., particularly Ms. Missy Tomlin, the Air Force Office of Scientific Research, and the Air Force Systems Command for making this Research Initiation Program effort possible. A special note of thanks is also due Mr. Jorge E. Beraun, Lt. Daniel M. Hodgkiss, and Mr. Dick Keppler of the Weapons Laboratory, Kirtland AFB, NM. Their essential support and helpful comments and guidance are greatly appreciated.

1.0 INTRODUCTION

The response of composite material pressurized cylinders to laser damage is a problem of considerable interest to the Air Force because of its relevance to Strategic Defense Initiative technologies. The problem arises out of the proposed use of lasers for strategic missile defense. The motorcases of many solid-fueled missiles are made of composite materials and it is this particular situation that the current problem addresses. Laser beams of various shapes and irradiances are of interest for such applications. The response of targets differs substantially depending on the particular material, structural, and laser parameters. Current analytical and computational models of this problem are largely empirically based and lack desired generality for predicting results over a wide range of size scales and new material systems of potential interest. The proposed research will address this need for more general and versatile models by seeking to incorporate more fundamental mechanisms and effects into the modeling.

The failure of a composite pressure vessel subjected to a deepening hole resulting from laser ablation is indeed complex. The complexities arise from a variety of sources, including the orthotropic nature of the material, a changing boundary condition with time (both thermal and mechanical), and material degradation due to extremely rapid temperature input. All of these complexities contribute to the final mechanical failure. It is unreasonable to initially approach a solution to this problem with all of the complexities included explicitly because excessive detail may in fact mask the primary sources of failure. A less complicated model, isolating particular mechanisms, will lead to a much better understanding of how a composite pressure vessel fails in addition to predicting a burst or a vent.

The primary concern of this research is understanding the mechanical failure process of a composite pressure vessel. For a pressurized composite cylinder, there are two final mechanical failure outcomes; the pressure vessel either bursts or vents. In a burst failure, the structural integrity of the pressure vessel is essentially lost,

whereas in the case of a vent, the structure is only damaged locally and depressurization occurs. Mechanical failures of laminated composite structures can be grouped into two general categories: in-plane failure and out-of-plane failure. In-plane failure is composed of matrix cracking and fiber breakage, whereas out-of-plane failure comes from essentially one source, delamination.

In many composite structures, including test article pressure vessels, these two failure modes are not necessarily independent of one another; that is, out-of-plane failure influences in-plane failure and vice versa. For this reason, it is important to develop an understanding of the mechanisms and sequences of interactions between in-plane and out-of-plane failure events. Previous efforts to model this problem (e.g. RDA, NRL, MDAC) have incorporated only in-plane failure mechanisms to analyze burst/vent behavior. Observed delaminatons in failed pressure vessel test specimens, however, indicate that the influences of out-of-plane failure cannot be neglected. Thus, a priority and major focus of this research has been to develop a model that incorporates the effects of delamination.

2.0 OBJECTIVES

The general aim of the Research Initiation Program effort has been to develop an improved computational model of the response of pressurized composite material cylinders to laser damage. In order to create a model more generally applicable to a variety of situations, the model should include more explicit physically-based representations of fundamental response and failure mechanisms. Specific features to be included in an improved model of this problem are delamination cracks and more accurate composite failure criteria.

In pursuing these objectives, the philosophy in the early stages has been to begin with a relatively simple model and implement improvements on this level first before addressing the more complex geometry of a pressurized composite cylinder with an external part-through hole. For an ideal "lower-order" model, the basic

mechanics of delamination initiation and growth as well as composite ply failure should be similar to that of the pressurized cylinder. The model can then be studied much more extensively and in greater detail in order to establish which features and effects should be of highest priority in further model development. The insight and knowledge gained from understanding the fundamental behavior of a simpler model will provide essential guidance for developing a more accurate yet computationally practical three-dimensional modeling strategy. The main emphasis initially is on delamination, or out-of-plane failure, and the specific goals can be summarized as follows:

(1) to identify a simplified model for studying the effects of delamination cracks in extensive detail and develop appropriate modeling strategies and procedures;

(2) to use the simplified model to gain a more complete understanding of delamination effects by generating a variety of specific solutions; and

(3) to incorporate in-plane failure and devise more complex models of the cylindrical pressure vessel configuration.

3.0 PRELIMINARY WORK

The intervening time between the conclusion of the 1989 Summer Faculty and Graduate Student Research Program and the beginning of the project period for the Research Initiation Program effort was spent primarily in literature review and further familiarization with hardware and software on the VAX and CRAY computers at Texas A&M University. The Composite Structures Modeling Workshop organized by Jorge Beraun of WL/TALE and Tom Edwards of RDA was also held during this same time period (5-6 December 1989). The main problem addressed by this meeting was the modeling and analysis of failure of the pressurized composite material test articles. The presentations and discussions proved extremely

useful in providing the latest information and findings, which helped further focus the initial direction of the research effort.

In order maintain a focus on the basic mechanics of this complex problem, the thermal degradation of material properties has been neglected. This is typically deemed acceptable because of the extremely high laser irradiances, in which case material removal is ablation dominated and occurs very rapidly. Under these conditons possible thermal degration of the material remaining around the laser hole is neglected because there is not enough time for significant heat transfer to occur. Therefore, laser action is simulated by successively removing material from the model. The basic mechanics of the problem can therefore be characterized as that of a composite pressure vessel with an external part-through hole.

As mentioned previously, the philosophy in the early stages was to begin with a relatively simple model and study this model first before addressing the more complex problem of a composite pressure vessel with an external part-through hole. The notched tensile bar [1], or NTB, was identified as a lower order analog of the pressurized cylinder configuration. It is essentially a two-dimensional specimen and can therefore be studied much more extensively and in greater detail. Further, the basic mechanics of delamination initiation and growth as well as composite ply failure are similar in many respects to that of the pressurized cylinder. In addition to understanding the effects of delamination on the failure of composite pressure vessels, it was desired develop procedures for modeling delamination growth explicitly and then characterize delamination behavior for a range of system parameters (e.g. delamination location and length, notch depth, material lay-ups, etc.).

Considerable early effort was directed at establishing procedures for constructing finite element models of this simplified system and obtaining solutions to each model. The PATRAN Plus (PDA Engineering) interactive graphics pre- and post-processor [2], which interfaces with a variety of analysis codes, was chosen for specifying details of the geometry and material properties of the finite element

models. PATRAN also provides a convenient graphical interface for displaying results in the form of deformed shapes and stress/strain contour plots. The finite element program that was decided upon for the actual analysis was ABAQUS by Hibbitt, Karlsson, and Sorensen [3]. A time-consuming but requisite task during this phase of the project was becoming familiar with many of the detailed commands and features for properly using these comprehensive software packages. Many different options dealing with boundary conditions, element types, solution procedures, and constraint equations were explored.

Explicitly modeling delamination cracks in the notched tensile bar specimen required the development of specialized procedures and programs for interfacing PATRAN data with ABAQUS. The commands available in ABAQUS for constraining nodes along a crack line and then releasing them as the crack grows are not directly supported by PATRAN. Thus, a customized program was developed to convert the PATRAN data file into a complete and correct ABAQUS input data file. Another issue that required considerable attention was the need to calculate strain energy release rates for delamination cracks. This kind of information is not standard output data for a typical finite element code. Displacements and forces at nodes in the crack tip can be used, however, to approximate strain energy release rates. In order to implement this so-called "virtual crack closure technique" another program was written for reading the required data from the ABAQUS output file and performing the appropriate calculations. Since the notched tensile bar produces a geometrically non-linear response, nodal forces and displacements had to first be transformed to a local coordinate system with the origin at the crack tip. This allows calculation of separate strain energy release rate components for modes I and II in addition to the total strain energy release rate. The computations were carried out using either a VAX 8800 mainframe or CRAY Y-MP2 supercomputer. PATRAN and ABAQUS are available on both machines. The calculations required to solve the finite element model and determine strain energy release rates for each case took about 5 minutes of CPU time on the CRAY or 1 to 1½ hours on the VAX.

4.0 NOTCHED TENSILE BAR STUDIES

The notched tensile bar with a delamination is shown schematically in Figure 1. The delamination initiates and grows from the notch region and is assumed to extend completely through the width of the specimen. Only half of the specimen is shown since the left end represents a plane of symmetry. The specimen is loaded with an axial tensile force (P). Since this analysis focuses on delamination behavior as a function of notch depth (s), delamination location (c), and delamination length, several models were investigated representing 3 combinations of these parameters.
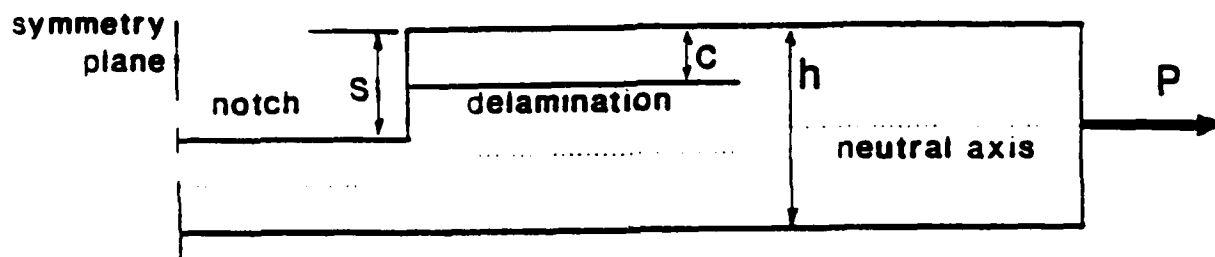
h = total specimen thickness
s = notch depth
c = delamination depth or location

Figure 1. Notched tensile bar geometry and nomenclature.

Particular questions that were addressed in studying the notched tensile bar are:

1. Are there characteristic trends exhibited by delaminations in a notched tensile bar test?
2. Where is delamination most likely to initiate?
3. Which fracture mode dominates delamination growth?
4. What is the effect of a deepening notch on delamination?
5. What is the effect of composite lay-up on delamination?

4.1 FINITE ELEMENT ANALYSIS

A typical finite element mesh for the NTB is presented in Figure 2. The geometry, meshing, material properties, loading and boundary conditions were created and specified using PATRAN Plus [2]. ABAQUS [3] was used to obtain solutions. Eight-noded isoparametric continuum elements with reduced integration (2x2) have been employed. In general, the finite element mesh consisted of approximately 2300 elements and 7400 nodes. Boundary conditions for the plane of symmetry in the notched region consisted of zero-valued displacements in the transverse direction, perpendicular to the axis of loading. Boundary conditions for the loaded end simulated a grip condition. Using the *EQUATION command in ABAQUS, all of the end nodes were constrained to have the same longitudinal

Figure 2. Finite element model of notched tensile bar.

displacement as the node at the specimen center line. The node at the center line was also constrained to have zero displacement in the transverse direction. Finally, a single concentrated load that had been normalized to the specimen width was applied to this node at the center line. In order to accurately treat the relatively large rotations and transverse displacements in the notched region, geometrically nonlinear analysis was required.

The three specific geometric configurations that were studied are summarized in Figure 3. Only the region in the vicinity of the delamination and the notch is

shown here. The delamination crack is centered in the refined portion of the mesh and the location is indicated by the arrow. The specimen was considered to be composed of a total of eight plies, each being 0.08 cm thick, for a total thickness of 0.64 cm. Both unidirectional, $4[0°]_s$, and cross-ply, $[0°/90°/0°/90°]_s$, lay-ups have been examined. The material was Kevlar/epoxy with the following properties:

$$E_1 = 80.6 \text{ GPa}, \ E_2 = E_3 = 13.7 \text{ GPa}, \ G_{12} = 2.04 \text{ GPa}$$
$$v_{12} = 0.343, \ v_{23} = 0.429, \ v_{31} = 0.043$$

The elements bounding the delamination crack were 0.022 cm x 0.020 cm in size, which is well within the recommendations of Raju, et al. [4] that crack tip elements be one-fourth to one-half of the ply thickness for total strain energy release rates to converge. The length of the refined mesh containing these elements was 1.32 cm, or approximately twice the total beam thickness. The range of delamination lengths was modeled by selectively treating coincident nodes along the line of the delamination as either connected to one another (uncracked portion) or separate and free to displace from one another (cracked portion). Plies far removed from the cracked region were modeled with one layer of elements per ply, as others have found adequate [5]. Intervening plies were modeled with two layers to provide a transition in meshing.

Strain energy release rates at the delamination crack tip were computed using the virtual crack closure technique (VCCT) [5]. Expressions for G_I and G_{II} for 8-noded isoparametric elements [7] require that displacements and forces at nodes near the crack tip be given in components tangent and normal to the plane of the crack. Thus, to accommodate the large displacements and rotations experienced by the NTB, a new local coordinate system was determined for each calculation of G. The tangential axis bisected the crack opening and was rotated from the longitudinal axis of the specimen. Forces and displacements were then transformed to the local coordinate system for calculating G_I and G_{II}.

Model A (s/h=c/h=0.125)

Uni	Cross
0°	0°
0°	90°
0°	0°
0°	90°
0°	90°
0°	0°
0°	90°
0°	0°

Model B (s/h=c/h=0.375)

Uni	Cross
0°	0°
0°	90°
0°	0°
0°	90°
0°	90°
0°	0°
0°	90°
0°	0°

Model C (s/h=0.500, c/h=0.375)

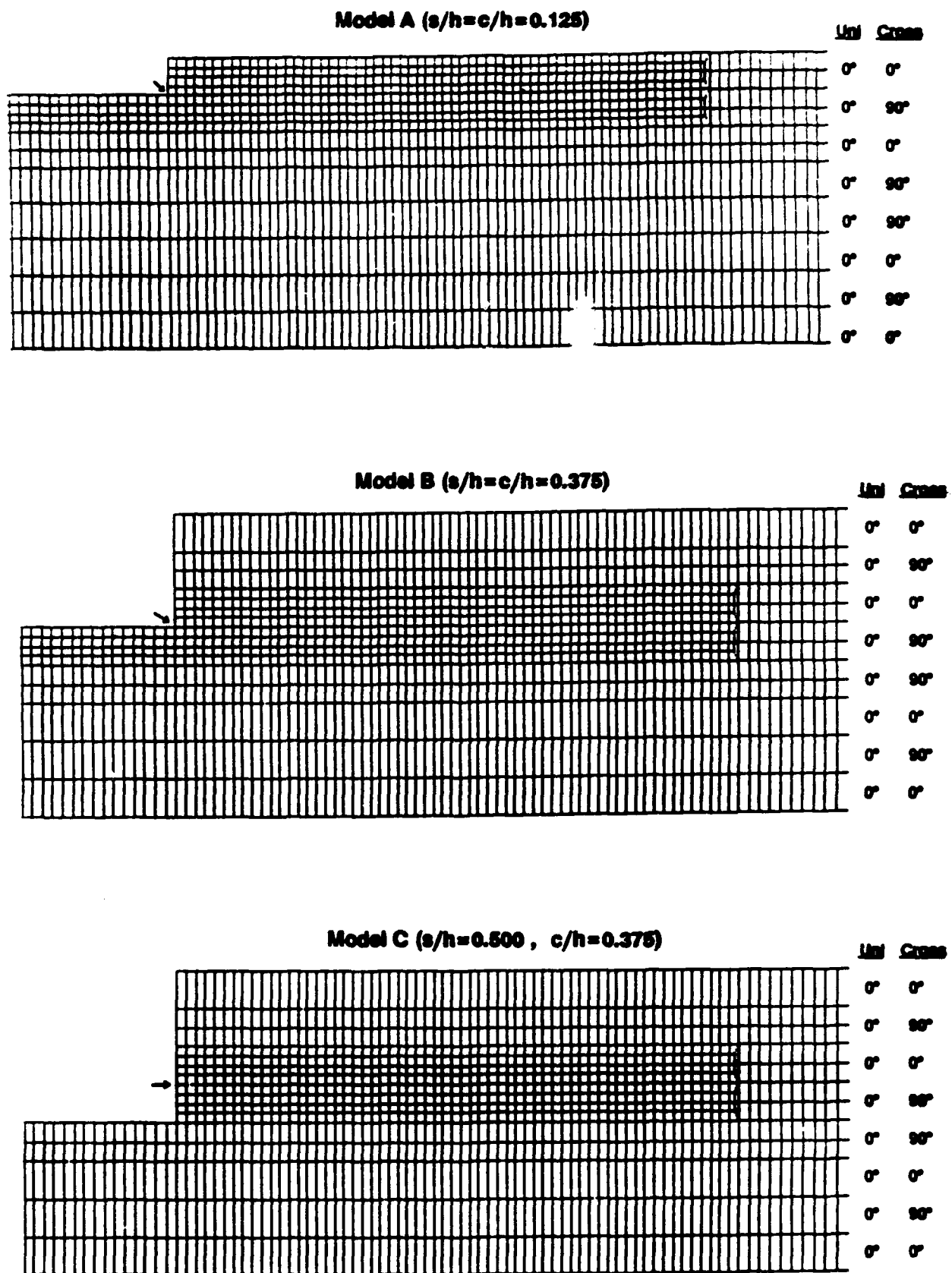| Uni	Cross
0°	0°
0°	90°
0°	0°
0°	90°
0°	90°
0°	0°
0°	90°
0°	0°

Figure 3. Specific models and lay-ups studied.

4.2 MODELING RESULTS

Each of the models was analyzed for delamination lengths of 0.022 cm, 0.33 cm, 0.66 cm, 0.99 cm and 1.32 cm. This procedure was applied to both the unidirectional and the cross-ply lay-up models. Representative results are summarized and briefly discussed below. More extensive results and discussions are available elsewhere [7].

4.2.1 Comparison with Analysis

An analytical solution for the available strain energy at delaminations in the NTB has been investigated by Simons for unidirectional laminates [1]. The solution uses a variational approach based upon simple beam-column assumptions for each of the three regions of the NTB with the Kirchhoff assumptions in effect, i.e. shear contributions were neglected.

Analytical and finite element solutions for the total strain energy release rate (G_T) as a function of delamination length and notch depth are shown in Figures 4 and 5. For a delamination at the bottom of the notch with the notch one ply thichkness deep (model A-uni; $s/h = c/h = 0.125$), the FE analysis and Simons' solution correspond fairly well. Although the differences are slight, the analytical results are consistently higher and decrease linearly, whereas the FE results show an initial increase followed by a decrease which parallels the analytical solution. These same two distinctive differences are seen more dramatically as the notch and delamination location deepen to three ply thicknesses (model B-uni; $s/h = c/h = 0.375$). For this case, the FE analysis is considerably lower. The behavior of the analytical solution shows a slightly non-linear decrease in G_T with delamination length, but the FE solution starts much lower, increases initially, and then again parallels the decrease of the analytical solution.

Figure 5 shows a comparison between the two methods for a delamination located one ply level above the bottom of the notch. Simons' solution not only.
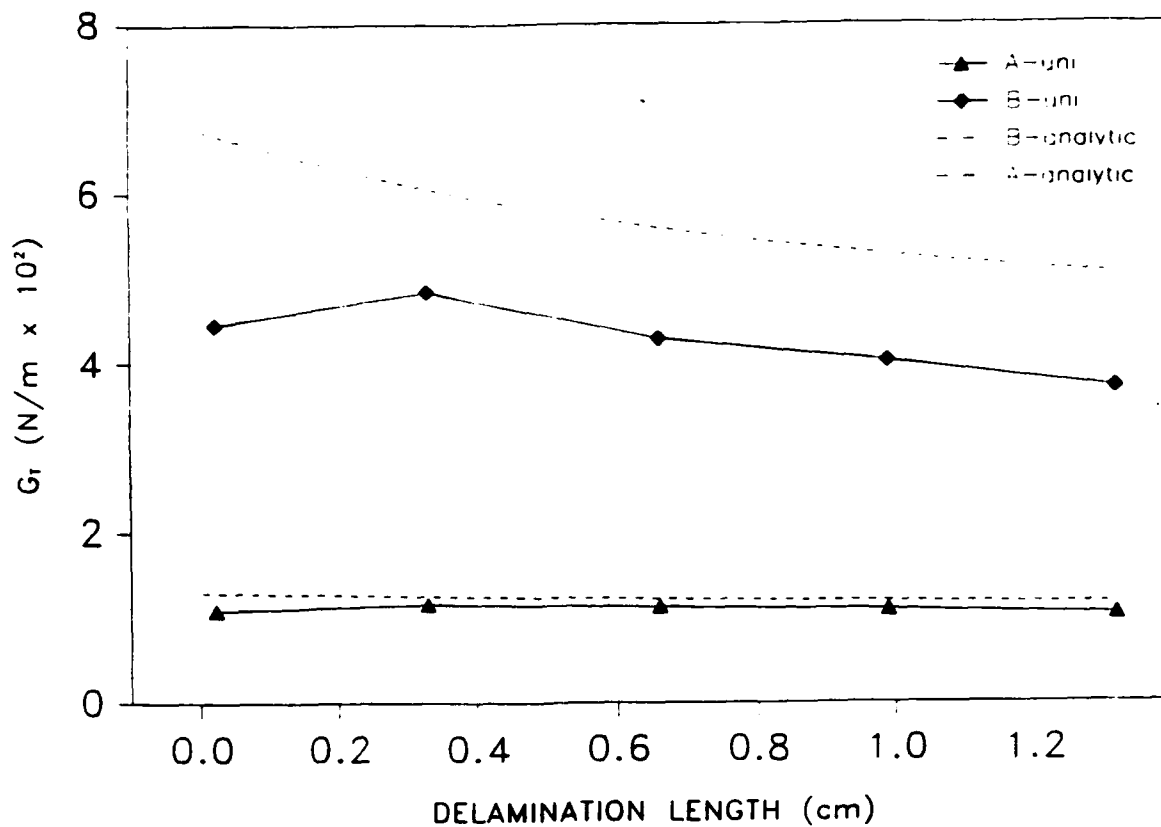
Figure 4. G_T for analysis and FE solutions with delamination at bottom of notch.
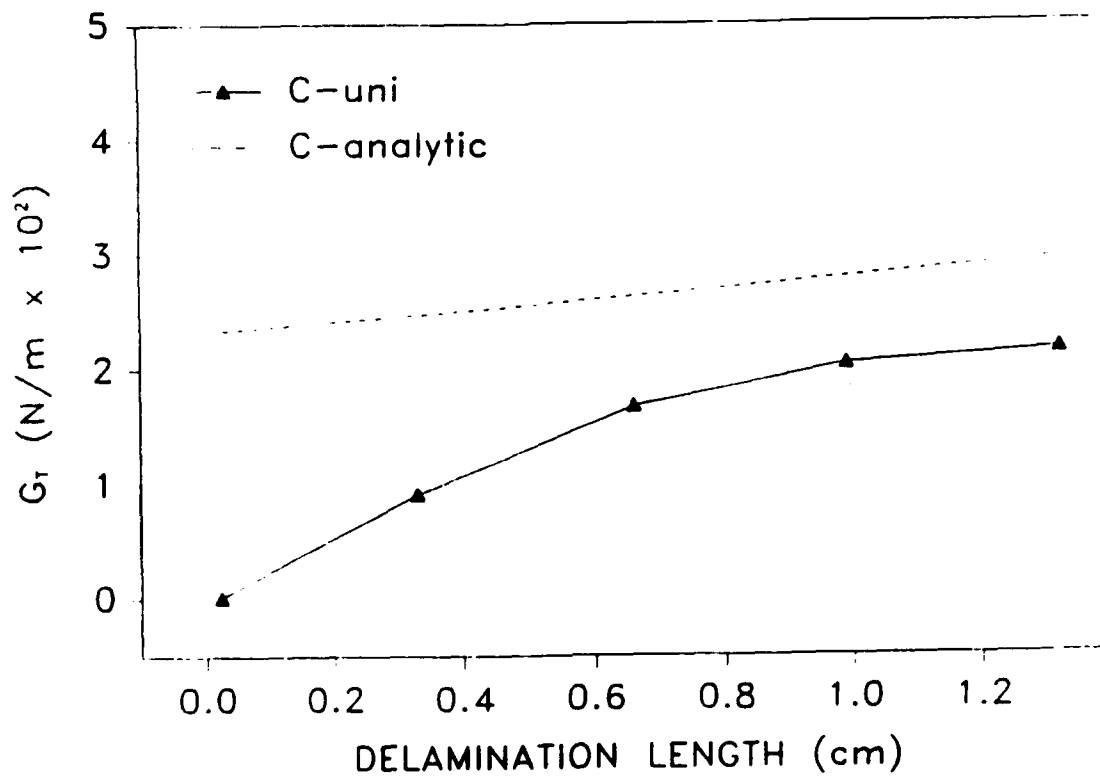

Figure 5. G_T for analysis and FE solution for delamination one ply from bottom of notch.

overestimates the available energy, but it predicts an essentially linear variation of G_T as a function of delamination length. The FE analysis predicts a very small amount of available energy supplied to the delamination initially, which then grows nonlinearly and begins to parallel the analytical solution

The generally lower G_T values for the FE models are probably best explained by considering the differences in the way shear deformation is treated in the two approaches. The analytical solution essentially neglects shear contributions to the strain energy, which results in stiffer solutions. The FE analysis, however, inherently includes shear deformation, which gives more compliant solutions and therefore lower strain energy release rates. These effects are most easily distinguished at the higher delamination lengths in Figures 4 and 5 where the two solutions parallel one another. Also note that the differences between the two solution methods is only slight for model A (Figure 4) because shear loading is relatively insignificant in this case. The shallow notch depth (one ply) results in very little bending and transverse displacement so the loading remains predominantly axial. As the notch depth increases, the eccentricity of the neutral axis increases causing more shear loading and more significant differences between the solutions.

While shear deformation provides a plausible explanation for the consistent overestimation of G_T by the analytical approach for longer crack lengths, discrepancies in the responses at short crack lengths are likely due to other effects as well. Recall that the analytical method treats the notched tensile bar as a series of beam-columns with the neutral axis offset to model the notch and delamination. Thus, at the vertical "wall" of the notch, the analytical approach predicts non-zero normal stresses since the bending moment and the axial force are both non-zero. The notch wall is actually traction free and the stresses should approach zero at the surface. The finite element method provides full-field stress distributions and therefore more accurately models the stresses in the vicinity of the notch wall. Accordingly, G_T from the FE solution starts from nearly zero for model C (Figure 5) while the analytical approach predicts a much larger G_T value for crack initiation.

In a similar manner, the FE solutions for models A and B in Figure 4 actually show a decrease in G_T as the crack length approaches zero.

4.2.2 Unidirectional Lay-up Models

The total strain energy release rates of the unidirectional lay-up models are shown in Figure 6. For delaminations located at the bottom of the notch (models A-uni and B-uni), the total available energy shows a nonlinear response as the delamination lengthens. Although the scale of the plot makes it less noticable for model A-uni, G_T increases initially and then steadily decreases with delamination length. This response suggests that once a delamination initiates it will initially grow rather readily, but instead of continuing in a potentially unstable mode the available energy decreases with crack growth beyond the early stages. The delamination may
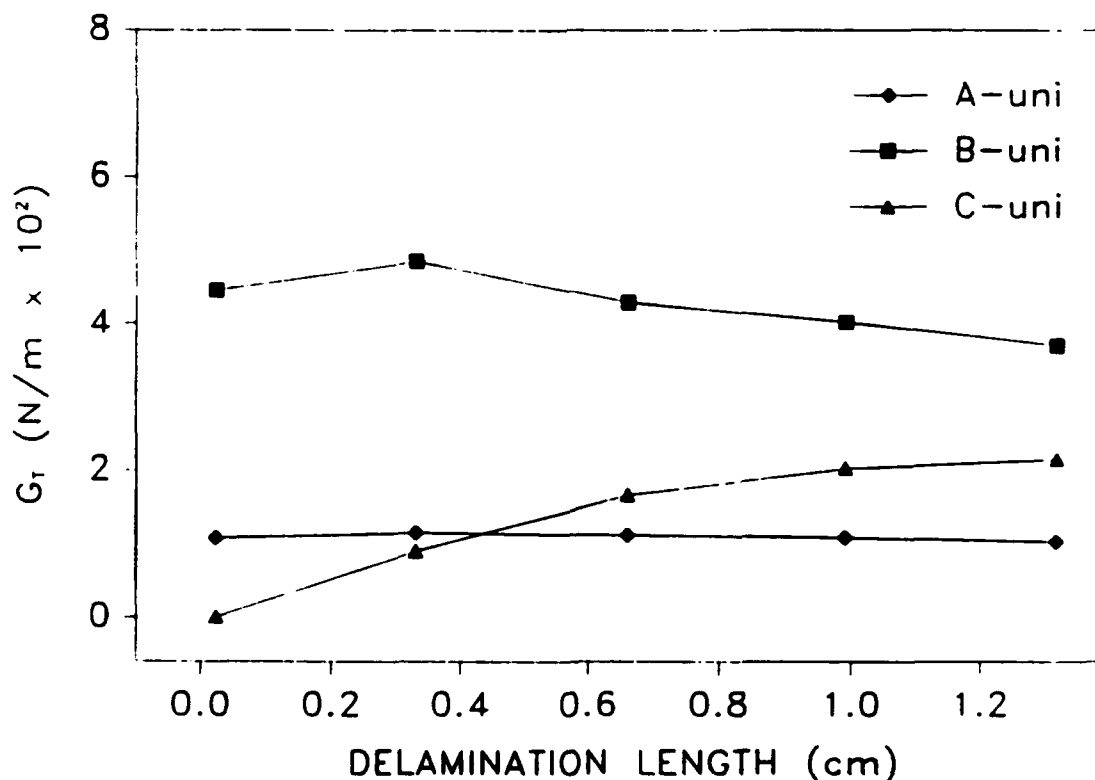

Figure 6. G_T for unidirectional models.

even arrest completely if G_T drops below the critical value. For delaminations located one ply above the bottom of the notch (model C-uni), the initial available energy is small, but shows a nonlinear increase as the delamination lengthens and begins to plateau. An increasing available energy suggests that if a delamination initiates and grows at a ply interface one ply from the bottom of the notch, it will grow unstably. With such an extremely low initiation energy, however, delaminations will probably not initiate at all unless they do so from the bottom of the notch.

For delaminations located at the bottom of the notch, models A-uni and B-uni indicate that the magnitude of the available energy for delamination initiation increases substantially with notch depth. In addition, models B-uni and C-uni provide a comparison of keeping the delamination depth fixed and increasing the notch depth one ply thickness. This comparison reveals a dramatic reduction of available energy for delamination initiation from 4.45 to 0.012 $\times 10^{-2}$N/m. As the crack grows, however, the strain energy release rates approach one another and become at least of the same order. An explanation for this behavior is that the corner of the notch creates a stress concentration and the locally intensified stresses contribute to higher available energies. The lower initiation energy in model C-uni is due to the wall (surface) of the notch being traction free and this region of the specimen being essentially stress free. As the crack lengthens, all of the material above the delamination becomes virtually free of load as well and acts as a new "effective" notch with a corner in the vicinity of the crack tip. The elevated stresses near this corner give rise to higher strain energy release rates.

A comparison of the total energy available for delamination growth for models B-uni and C-uni also indicates stable delamination behavior under the scenario of a steadily deepening notch. If a delamination initiates and grows to say 0.33 cm in model B-uni, and then the notch is allowed to deepen by one ply such that model C-uni is obtained, the total available energy supplied to the delamination drops, which may result in delamination growth termination. This trend is seen for all of the crack lengths studied, indicating that notch deepening would tend to dissuade delamination

growth. As a practical consequence, this behavior suggests that the magnitudes of the relative rate of delamination crack growth and the rate of notch deepening will play an inportant role in determining how much delamination actually occurs.

Similar plots of total strain energy release rates (G_T), and mode I and mode II components (G_I & G_{II}, respectively), clearly demonstrate a mode II dominance in the delamination response [8]. This is consistent with the mode II dominance reported by Hooper and Hwu [9] for the cracked lap shear specimen, which is similar to the NTB. The only significant departure from mode II dominance occurs for very short delamination lengths for models A and B.

4.2.3 Cross-ply Lay-up Models

The total strain energy release rate (G_T) for the three cross-ply lay-up models is shown in Figure 7. The general trends are similar to the unidirectional models.
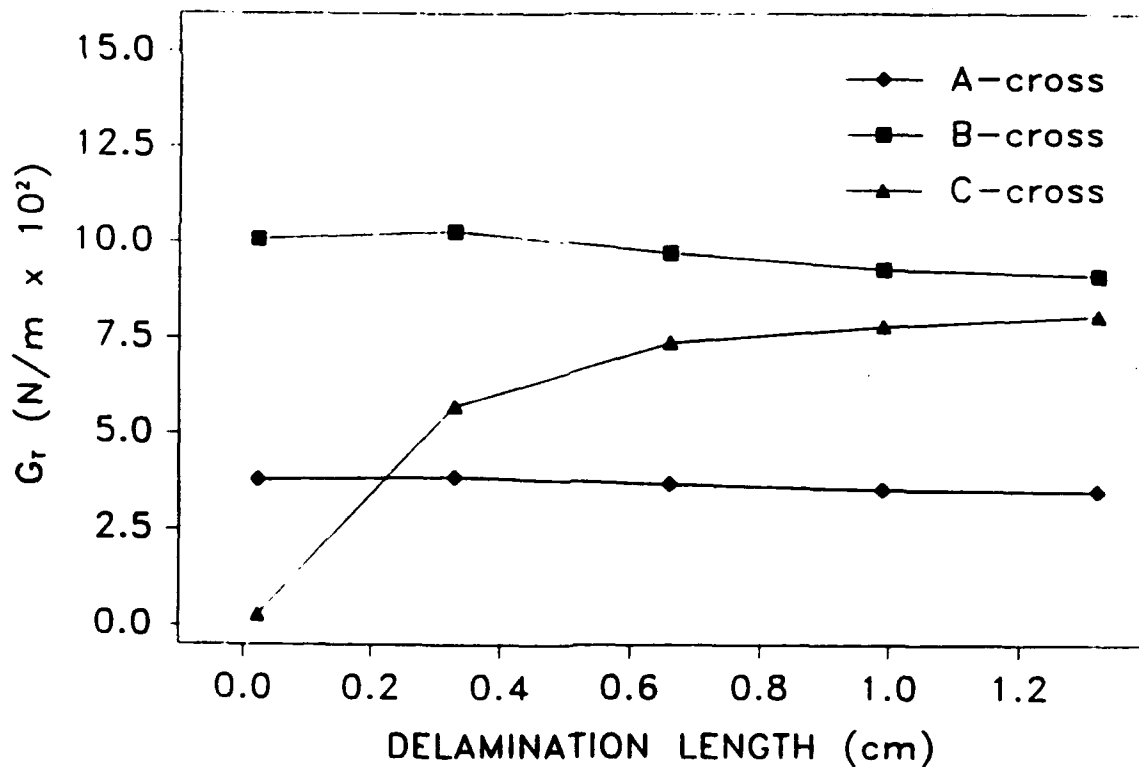

Figure 7. G_T for cross-ply models.

The notable exception is that the G_T values for the cross-ply cases are consistently higher than for corresponding unidirectional cases. As with the unidirectional models, the cross-ply models A-cross and B-cross show that the intensity of the available energy for delamination intitiation increases with notch depth, indicating a higher likelihood of delamination initiation. G_T also increases slightly as the delamination initiates and then decreases steadily with crack length as was observed for the unidirectional models (A-uni and B-uni). Comparing models B-cross and C-cross shows again the rather dramatic drop in crack initiation energy when the notch is deepened by one ply thickness and the delamination remains at the same location. The G_T values also approach one another as the crack lengthens suggesting that, as before with the unidirectional lay-ups, if delamination does not initiate at the bottom of the notch then it is not likely to initiate subsequently at the same location.

Considering the individual components of the energy release rates, the dominant mode for delamination initiation and growth is again the shearing or sliding mode, mode II. These results [8] also provide some insight into the generally higher G_T values for the cross-ply models compared to the unidirectional models. The delamination in the cross-ply models is located between a $0°$-ply and a $90°$-ply. The substantial mismatch in stiffness and stresses resulting from this gives rise to higher G_{II} values while G_I remains roughly the same. The mode II dominance of the cross-ply models is therefore much greater than for corresponding unidirectional models.

4.3 SUMMARY COMMENTS

Several specific conclusions can be drawn from the finite element modeling of the NTB. Those that apply to both unidirectional and cross-ply lay-ups of the notched tensile bar are:

- G_T exhibits a slight increase followed by a steady decrease with delamination length for delaminations at the notch bottom.

- G_T exhibits a nonlinear increase with delamination length for delaminations one ply above the notch bottom.

- Delamination is most likely to initiate at the notch bottom.

- G_T increases with notch depth for delaminations at the notch bottom.

- Mode II dominates delamination growth.

Conclusions drawn from comparing cross-ply and unidirectional models are:

- G_T values are generally significantly higher for cross-ply lay-ups compared to corresponding unidirectional cases.

- Mode II dominance is greater for cross-ply lay-ups.

The practical implication of these findings is that delaminations are most likely to initiate from the notch bottom and grow in a stable manner until the crack arrests, or stops growing. Furthermore, once the external hole grows past a current delamination, the crack is not likely to grow any further or re-initiate.

The results to date clearly demonstrate the value of studying a simplified model to understand basic trends in the mechanics of this type of problem. The level of detail that is included in the two-dimensional finite element model of the notched tensile bar would be prohibitive in a fully three-dimensional treatment of the cylindrical bottle geometry. A tangible benefit that has already been realized

is that delamination initiation behavior only needs to be considered for cracks emanating from the bottom of the external hole. This begins to limit the range of system parameters that needs to be incorporated into the more complicated model. As a result, development of the three-dimensional model will be more efficient and can be focused on the most essential features and phenomena. In addition, on-going studies of the notched tensile bar are examining the effects of curvature on such models [10]. Results from these studies will similarly aid in three-dimensional model development.

5.0 DISSEMINATION OF RESULTS

Preliminary results and findings have been disseminated in several forms and on several occasions since the end of the 1989 Summer Faculty and Graduate Student Research Program. Detailed citations are given below. The major conclusions and basic research strategy from the work conducted during the summer of 1989 were summarized in an oral presentation at the Composite Structures Modeling Workshop, December 5-6, 1989, Albuquerque, NM, which was jointly sponsored by WL/TALE and R & D Associates. Monthly activity reports have been submitted to WL/TALE since the beginning of the formal grant period (January 1990). These have typically taken the form of 1-2 page memos. A more comprehensive Progress Report accompanied by a Research Proposal were also submitted to WL/TALE in October of 1990.

More formal and general technical presentations have also resulted from the detailed modeling studies of the notched tensile bar. The first of these was an oral slide presentation at the Society for Experimental Mechanics Spring Conference held June 4-6, 1990, Albuquerque, NM. This trip also provided an opportunity to present an informal briefing to Mr. Jorge E. Beraun of the Weapons Laboratory. In addition, a full-length paper for the 1991 ASME Energy Technology Conference and Exhibition has recently been accepted for

presentation and publication. Finally, a paper and presentation for the 1991 AIAA SDM Conference has also been accepted.

Citations

Hogan, H. A., "Improved Modeling of the Stressed Composite Cylinder to Laser Damage," presented at the Composite Structures Modeling Workshop jointly sponsored by the United States Air Force Weapons Laboratory Technology Assessment Office and R&D Associates, 5 December 1989, Albuquerque, NM.

Harbert, S. J. and Hogan, H. A., "An Investigation of the Effects of Delamination on the Response of Pressurized Composite Cylinders," invited presentation, 1990 Society for Experimental Mechanics Spring Conference, June 4-6, 1990, Albuquerque, NM.

Harbert, S. J. and Hogan, H. A., "An Analysis of Delamination in a Fiber Reinforced Composite Material Notched Tensile Bar," accepted for presentation and publication, Composite Material Symposium, ASME Energy Technology Conference and Exhibition, Jan. 20-24, 1991, Houston, TX.

Harbert, S. J. and Hogan, H. A., "Strain Energy Release Rates in Straight and Curved Notched Composite Beams," accepted for presentation and publication, AIAA 32nd Structures, Structural Dynamics, and Materials (SDM) Conference, April 8-10, 1991, Baltimore, MD.

6.0 RECOMMENDATIONS

Based upon the findings from the research conducted thus far, the first priority for further research is to continue studying the notched tensile bar specimen, since it can be treated in great detail and examined more comprehensively. Using this model, the emphasis should be on characterizing the detailed interactions and sequences of events when both in-plane (ply failure) and out-of-plane (delamination) failure modes are occurring essentially simultaneously. The special procedures for using PATRAN and ABAQUS to conduct this kind of analysis need to be further improved and automated to require a minimal amount of direct "hands on" user interaction. Simulations can then be conducted for various combinations of material properties, such as critical strain energy release rates (which govern delamination crack growth) and single ply ultimate strains (which govern ply failure).

Another major emphasis in studying the notched tensile bar should be to investigate ways to use beam elements to generate lower-order models of the system. The current effort utilizes two-dimensional continuum finite elements and extending this level of detail to the cylindrical pressure vessel configuration would require three-dimensional continuum (solid "brick") elements, which would be too computation intensive to be of much practical value. Thus, a lower-order model is essential for ultimately simulating failures of the composite test bottles. The critical challenge in developing such models, however, is to balance model simplicity and practicality with accuracy and fidelity. In order to be useful as a predictor and evaluation tool, the model must at least capture the essential elements of the mechanical response and failure processes. One of the main benefits of the extensive two-dimensional modeling of the notched tensile bar is to provide quantitative evidence and guidelines on which features of the response are absolutely necessary in the three-dimensional model and which are of secondary importance.

As lower-order modeling schemes are evaluated for the notched tensile bar configuration, similar approaches should be developed and implemented for the 3-

dimensional geometry of the cylindrical pressure vessel. Options to be considered will range from using "thick" shell elements, which include shear deformation, to stacking layers of thin shell elements. The stacked shell element approach could be as simple as treating each ply above and below delaminations as a single layer of shell elements or it could involve explicitly treating each individual ply as a separate layer of shell elements. Any scheme involving stacked shell elements will require considerable development effort in order to devise a method for appropriately constraining the element degrees of freedom through the thickness. Another approach that might be examined would be to use a single layer of shell elements but to use elements that have the effects of embedded delamination cracks included in the element constitutive relations. These typically involve some sort of internal state variable, which effectively introduces another unknown material property, but the computational requirements are reduced significantly.

Ultimately, the most promising modeling strategy should be selected in order to conduct an extensive series of simulations. The goal of these simulations would be to model the essential elements of the failure processes (both in-plane and out-of-plane) and to ultimately pre :t the final failure condition and mode. A successful lower-order model of the pressurized composite cylinder configuration will allow investigation of a variety of system parameters. The effects of such variables as composite lay-ups, stacking sequences, and material systems can then be characterized theoretically. Different size scales, diameters, and thicknesses could also be studied.

REFERENCES

1. Simons, D., "Delamination and Nonlinear Geometric Effects in Laser-Induced Composite Failure," presented at Composite Structures Modeling Workshop sponsored by Air Force Weapons Laboratory (WL/TALE) and R & D Associates, 5-6 December 1989.

2. PATRAN Plus User Manual, PDA Engineering, Costa Mesa, California, 1989.

3. ABAQUS User's Manual, Hibbitt, Karlsson & Sorensen, Inc., Providence, Rhode Island, 1989.

4. Raju, I. S., Crews, Jr., J. H., and Aminpour, M. A., "Convergence of Strain Energy Release Rate Components for Edge-Delaminated Composite Laminates," NASA Technical Memorandum 89135, April 1987.

5. Johnson, W. S. and Mall, S., "A Fracture Mechanics Approach for Designing Adhesively Bonded Joints," Delamination and Debonding of Materials, ASTM STP 876, W. S. Johnston, ed., ASTM, Philadelphia, 1985, pp. 189-199.

6. Rybicki, E. F. and Kanninen, M. F., "A Finite Element Calculation of Stress Intensity Factors by Modified Crack Closure Integral," Engineering Fracture Mechanics, Vol. 9, 1977, pp. 931-938.

7. Raju, I. S., "Simple Formulas for Strain-Energy Release Rates with Higher Order and Singular Finite Elements," NASA Contractor Report 178186, December 1986.

8. Harbert, S. J. and Hogan, H. A., "An Analysis of Delamination in a Fiber Reinforced Composite Material Notched Tensile Bar," accepted for presentation and publication, Composite Material Symposium, ASME Energy Technology Conference and Exhibition, Jan. 20-24, 1991, Houston, TX.

9. Hooper, S. J. and Hwu, C., "A Mixed Mode Fracture Analysis of the Crack Lap Shear Specimen Using a Sublaminate Approach," AIAA 30th Structures, Structural Dynamics and Materials Conference, Mobile, Alabama, April 3-5, 1989, paper no. 89-1255-CP, pp. 830-836.

10. Harbert, S. J. and Hogan, H. A., "Strain Energy Release Rates in Straight and Curved Notched Composite Beams," accepted for presentation and publication, AIAA 32nd Structures, Structural Dynamics, and Materials (SDM) Conference, April 8-10, 1991, Baltimore, MD.

1988 USAF-UES RESEARCH INITIATION PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.

FINAL REPORT

RELATIVISTIC EFFECTS IN GLOBAL POSITIONING

Prepared by: Arkady Kheyfets, Ph. D.

Academic Rank: Assistant Professor

Department and   Department of Mathematics
University: North Carolina State University
                    Raleigh, NC 27695-8205

USAF Researcher: Dr. Warner A. Miller
                    WL/AWPP
                    Kirtland AFB, NM 87117-6008

RELATIVISTIC EFFECTS IN GLOBAL POSITIONING

by

Arkady Kheyfets

ABSTRACT

Precise global clock synchronization is an integral part of the Global Positioning System (GPS) operations. The GPS satellites clocks are moving with respect to the clocks of the surface stations observers at speeds sufficient to necessitate careful consideration of special relativistic effects on synchronization of the clocks. At the same time, the orbits radii of the GPS satellites are large enough to cause a difference between the gravitational potentials at the satellites clocks and at the surface stations clocks sufficient to produce effects on the clock synchronization of the same order of magnitude as the special relativistic effects. A consistent treatment of both effects can be done only in general relativity.

We have performed a general relativistic analysis of the GPS time transfer effects and the ranging procedure in GPS using both traditional mathematical techniques and the newly developed general relativistic technique of null strut calculus. The expressions for the effects, which were obtained, admit unambiguous physical interpretation of each term, which clarifies the physical origin of the effects. We give a systematic and and a complete analysis of the relativity produced errors in the GPS ranging. Our results provide a firm theoretical basis for making estimates of the accuracy of the existing GPS procedures as well as evaluating any future changes in it. The null-strut calculus technique illuminates the 4-geometry of the procedure. The null-strut calculus looks very promising as a future common language in the formulation of such problems and the development of a satellite-based Spacetime Common Grid (SCG).

ACKNOWLEDGMENTS

I. INTRODUCTION.

Proper operating of the satellite–based Global Positioning System (GPS) imposes increasingly demanding requirements on global clock synchronization. Such synchronization account for both special relativistic effects and the effects caused by the earth's gravi .tional field. A consistent treatment of both kinds of effects can be performed only within the framework of general relativity.

Almost all previcus attempts to resolve the problem were undertaken within the framework of special relativity. Furthermore, there was considerable difference in the results obtained by different researchers evaluating the special relativistic effects, and some apparent confusion concerning formulation of the problem.

The 4–Geodesy Section of the Advanced Concept Branch of the USAF Advanced Weapons Laboratory at Kirtland Air Force Base recognized that general relativistic solution of the problem was a necessity. It was also recognized that the language of 4–geometry and, in particular, the newly developed null strut calculus could be most helpful in formulating the problem.

My research interests have been in the area of applications of modern mathematical methods in field theories, foundational problems of physics, and, especially, classical and quantum gravity. I have a very strong background in general relativity, both in foundational aspects and in applications. I owe it to the University of Texas at Austin (where I received my Ph. D. degree), and, in particular, to Prof. John A. Wheeler (with whom I worked for more than four years). My particular strength is the ability to see clearly the geometric content of general relativistic problems. This geometric insight played a very important role in performing this research and, no doubt, will be used effectively in the future.

II. OBJECTIVES OF THE RESEARCH EFFORT

The requirements on global clock synchronization are becoming increasingly demanding in GPS operations. Furthermore, when more satellites are added to the GPS constellation to form a spacetime common grid (especially with cross-link ranging between satellites), it is believed that the precision requirements will become crucial for the coherent functioning of the system as a whole.

The GPS constellation will someday contain 18 clocks (with 3 active spares) moving with respect to each other. The GPS satellites will have almost circular orbits of 4 earth radii, with 12-hour periods, which means that the satellite velocity will be ~ 8 times the velocity of the surface station observer originating from earth rotation. Therefore, the required precision of the clocks synchronization necessitates taking into account special relativistic effects on the rate of the clocks. In addition, all of the activity of the GPS constellation occurs in the earth's gravitational field with clocks placed in positions with different gravitational potentials. The gravitational influence of the field on a clock's rate is determined by the parameter $\frac{Ma}{r}$ which produces an effect of the same order of magnitude as the second order special relativistic effects. A consistent treatment of both effects together can be done only within the framework of general relativity.

The need for a general relativistic treatment of the GPS clock synchronization problem was recognized prior to the 1988 SFRP[1] and confirmed during it[2]. Investigation of the general relativistic effects on clock rates was performed on several occasions[16]. The results have been implemented partially into the ranging procedure. However, by the time the 1988 SFRP period started the matter became a subject of controversy[3]. Previous results were ignored (to such an extent that we became informed about them only by the end of the 1988 SFRP period). Everything was started anew. Two schemes, one using the second order Doppler correction[4,5], and another one using the ranging data[2,6], were under consideration. There was

considerable difference in the results of the calculations of the second order Doppler correction. Most of the new attempts[4,6] were undertaken within the framework of special relativity (with different results) with only one completely general relativistic attempt[5]. Almost all the results did not look intuitively satisfactory. Some of them contained nonlocal terms having no physical interpretation. Others looked suspiciously symmetric with respect to the transmitter and the receiver, even though the expression for the Doppler shift itself was not, and one would expect this asymmetry to show up in the second order approximation. Only the results of Ashby[4] looked perfect in all intuitive respects but, unfortunately, they were special relativistic. The overall impression was that different authors did not always quite understand each other and, possibly, had tried to calculate different things.

Working within the 1988 SFRP (May – July, 1988) I used the technique of the tensor series expansion (1) to find the correct general relativistic expression for the Doppler shift in the earth's gravity field up to the second order and to give the results physical interpretation, and (2) to formulate and solve the problem of global clock synchronization in the earth's gravitational field.

In my work, I utilized maximally the 4–geometric language in formulating the problems and pictorial demonstrations of the problems peculiarities, having as a goal to avoid confusion in interpretation of the results in the future.

The analysis performed by me during the 1988 SFRP (cf. my 1988 SFRP Final Report[17]) was done only for the simplest cases. For instance, I restricted my analysis to the case of circular orbits when comparing the rates of the clock of the ground observer and the satellite clock, and to the case the equatorial ground observer and equatorial plane of the satellite orbit when considering the initial clock synchronization (in mathematical language, when evaluating the constants of integration). The simplifications were necessary to stress the physics of the problems and to get rid of the details that did not have the relativistic origin. They allowed me to achieve a clear understanding of the key relativistic features involved in GPS

time transfer.

The main results of the 1988 SFRP research are used extensively in sections III, and IV of this report. Section V analyses a misconception that lead to the controversy mentioned above. It is done not because the controversy was really important, but to demonstrate the advantages of the geometric approach to the relativistic GPS problems.

My objectives for the research efforts within the RIP were dictated by the necessity to learn how the basic relativistic effects described previously in the simplest situations look in more realistic setting:

(1) Although the orbits of the GPS satellites ideally should be circular, it is clear that they cannot be perfectly circular. Consequently, the question is how the expressions comparing the rates of clocks must be modified if the orbits are slightly noncircular (if they indeed should be modified).

(2) The initial clock synchronization procedure was considered by us previously for the highly idealized case of the equatorially placed ground observer and equatorial satellite orbits. It was shown that the procedure allowed to initially set the clocks in such a way that the constant of integration in the formula relating the ground observer and the satellite clock time became equal to zero with a nonaccumulating error of the order of $M_\oplus/r$ (cf. section IV of this report). However the GPS satellite orbits planes are inclined with respect to the equatorial plane. Also, the ground observer ethalon clock is not placed on the equator. The question is whether the initial synchronization procedure can be made to work with the same precision in this more realistic setting.

(3) The basic relation between the range data and time data standardly used in GPS is the special relativistic expression

$$d_{AB} = c(t_B - t_A),$$

where c is the speed of light, or, in the system of units with $c = 1$ (commonly accepted in relativity [8,9,17])

$$d_{AB} = t_B - t_A$$

Here A and B are two events in spacetime related by a light signal (or, in mathematical language, connected by a null geodesic), d_{AB} is the range between A and B, and t_A, t_B are the times of events A and B, respectively. The range d_{AB} and the times t_A, t_B in special relativity are measured in the frame of the same (but arbitrary) global inertial observer, or, to put it in different language, in the same global orthonormal coordinate system in spacetime. Neither the concept of a global inertial frame nor the concept of a global orthonormal coordinate system makes sense in in curved spacetime for the general gravitational field. This makes the interpretation of the range–time–data relation ambiguous in the general situation. However in the case of the weak, static, spherically symmetric gravitational field (which is the case in all GPS problems) one might hope to find a global interpretation of the relation. The relation then will become an approximate one, and the question is whether it can be kept precise enough to meet the requirements of the GPS.

In sections V–VIII we analyze systematically the relativity produced errors in ranging. This analysis has never been done before. Our results provide a firm theoretical basis for making estimates of the accuracy of the existing GPS procedures as well as evaluating any future changes in it.

III. DOPPLER SHIFT IN A SCHWARZSCHILD FIELD.

We use the Schwarzschild geometry as the model of the earth's gravitational field. In doing, so we neglect contribution of the earth rotation in the gravitational field. An enhanced model would involve the Kerr metric. However, the evaluation of the Kerr model parameters shows that the produced effect of the inertial frames dragging would be of higher order than the effects caused by the parameters coming from the Schwarzschild model[7]. Meanwhile, the estimate of the Schwarzschild model parameters shows that the effect of $M_{\oplus}/r$ (here $M_{\oplus}$ is the earth mass and r is the Schwarzschild radial coordinate[8,9]) is of the same order as effects of the squares of the relevant velocities. This is obvious for the satellite orbits. Indeed, the Kepler's law for circular orbits (and it is well known that the Kepler's law is satisfied exactly for circular orbits even in general relativity[8]) reads $M_{\oplus} = v^2 r$ so that $M_{\oplus}/r = v^2$. For the ground observer v^2 is less than $M_{\oplus}/r$. Nevertheless, for all the situations considered in GPS one can consider that $v^2 \sim M_{\oplus}/r$.

Thus, to discuss Doppler shift up to the second order, we can use Schwarzschild geometry as a model of the earth's gravitational field and, provided that in all approximations terms proportional to $M_{\oplus}/r$ are retained (we can neglect higher powers of $M_{\oplus}/r$), we obtain a satisfactory expression for the Doppler correction up to the second order with respect to the velocities involved in the picture. We can also say that all the relations below are satisfied up to the second order with respect to v or $\sqrt{M_{\oplus}/r}$. In general, throughout this report the abbreviated expression "the relation is satisfied up to the nth order" means that it is satisfied up to the nth order in v and $\sqrt{M_{\oplus}/r}$.

The Schwarzschild geometry is a static spherically symmetric geometry. Its metric in Schwarzschild coordinates is given by the expression[8]

$$ds^2 = -\left(1 - \frac{2M_{\oplus}}{r}\right) dt^2 + \left(1 - \frac{2M_{\oplus}}{r}\right)^{-1} dr^2 + r^2 \left(d\theta^2 + \sin\theta d\phi^2\right), \quad (1)$$

where t, r, θ, and ϕ are Schwarzschild coordinates*, and $M_\oplus$ is the mass of the earth.

In the geometric picture (cf. Fig. 1) describing the Doppler shift of an electromagnetic signal sent from the transmitting satellite to the surface station observer, the free falling satellite has a geodesic world line, whereas the observer, being attached to the earth, has a world line with all three of the curvatures nonzero[9]. A Doppler shift arises since the 4–velocity of the satellite (at the moment of signal transmission) and the observer (at the moment of receiving) are not parallel. More precisely, the result of the parallel transport of the satellite 4–velocity along the null geodesic connecting the event of transmitting and the event of receiving does not coincide with the 4–velocity of the observer.

The frequency shift can be expressed in terms of the 4–velocities of the satellite and the observer and the 4–momentum of the photon traveling from the satellite to the observer

$$\mathcal{D} = \frac{\nu_S - \nu_O}{\nu_S} = \frac{{}_SP_\mu V_S^\mu - {}_OP_\mu V_O^\mu}{{}_SP_\mu V_S^\mu}, \tag{2}$$

where V_S^μ, V_O^μ are the 4–velocities of the satellite and the observer and ${}_SP_\mu$, ${}_OP_\mu$ are the photon 4–momentum at the event of transmitting and the event of receiving, respectively. The 4–momentum of the photon is parallel transported along the null geodesic connecting the events of transmitting and receiving and is tangent to the null geodesic at all times.

We have used for calculation of the Doppler shift (up to the second order) the technique of the tensor series expansion of the world function[9] developed by J. L. Synge. Here we only describe and explain the result. Following J. L. Synge we introduce new coordinates $(x^\mu)_{\mu=0,1,2,3}$ related to the Schwarzschild coordinates as

---

* We use throughout this report the system of units commonly accepted in general relativity with both the velocity of light and the gravitational constant equal to unity
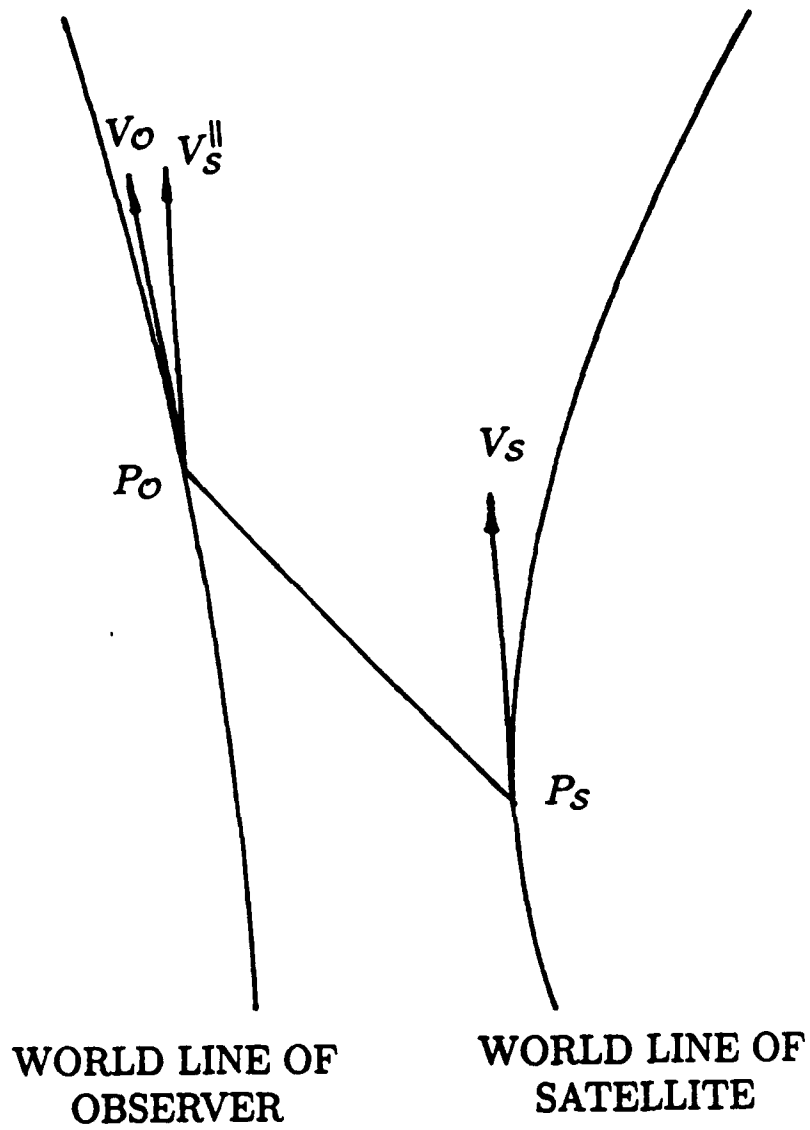
Fig. 1. Geometry of the Doppler shift. The vector $V_S^{\parallel}$ is the result of the parallel transport of V_S along the null geodesic $P_S P_O$. The Doppler shift is caused by $V_O \neq V_S^{\parallel}$.

follows

$$x^0 = t, \quad x^1 = r \sin \theta \cos \phi, \quad x^2 = r \sin \theta \sin \phi, \quad x^3 = r \cos \theta. \tag{3}$$

The metric tensor in these coordinates can be expressed as the sum $g_{\mu\nu} = \eta_{\mu\nu} + \gamma_{\mu\nu}$, where $\eta_{\mu\nu} = \text{diag}(-1,1,1,1)$ and $\gamma_{\mu\nu}$ are small and static ($\gamma_{\mu\nu,0} = 0$). The coordinates (x^μ) are very convenient for a pictorial representation of the Doppler shift. In Fig. 2 these coordinates are used as coordinates of a Euclidean space. Of course, in this space the geodesics of the original Schwarzschild space do not always look like straight lines. The world line of the satellite in this picture is geodesic but looks curved. The vertical straight lines are the integral lines of the timelike Killing vector field of the Schwarzschild metric (described by the equations $x^i = \text{const}, \; i = 1, 2, 3$). The satellite and the observer are moving with respect to Schwarzschild coordinates, so that the 4-velocities V_S, V_O are not parallel to the Killing vectors $\frac{\partial}{\partial t} = \frac{\partial}{\partial x^0}$. The angles between V_S, V_O and the Killing vectors $\frac{\partial}{\partial t}$ (directed upward) are different and determined by the satellite and the observer orbital velocities. If the satellite and the observer were at rest with respect to Schwarzschild coordinates (in which case their world lines would be pictured as vertical straight lines), we would get for the Doppler shift

$$\mathcal{D} = \frac{M_\otimes}{R_S} - \frac{M_\otimes}{R_O}. \tag{4}$$

The right hand side of Eqn. (4) is often called the gravitational Doppler shift. It is of second order in magnitude and should be expected to appear as one of the terms in the final result.

In fact, the final result for the Doppler shift up to the second order is

$$\mathcal{D} = (V_O^i - V_S^i)\frac{\Delta x^i}{\Delta t} + (V_O^i - V_S^i)\frac{\Delta x^i}{\Delta t} V_S^k \frac{\Delta x^k}{\Delta t} + \left(\frac{M_\otimes}{R_S} - \frac{M_\otimes}{R_O}\right) + \frac{1}{2}(V_S^i V_S^i - V_O^i V_O^i) \tag{5}$$

where $i, k = 1, 2, 3$, $\Delta x^i = x_O^i - x_S^i$, $\Delta t = x_O^0 - x_S^0$, and the summation over repeating indices is assumed.
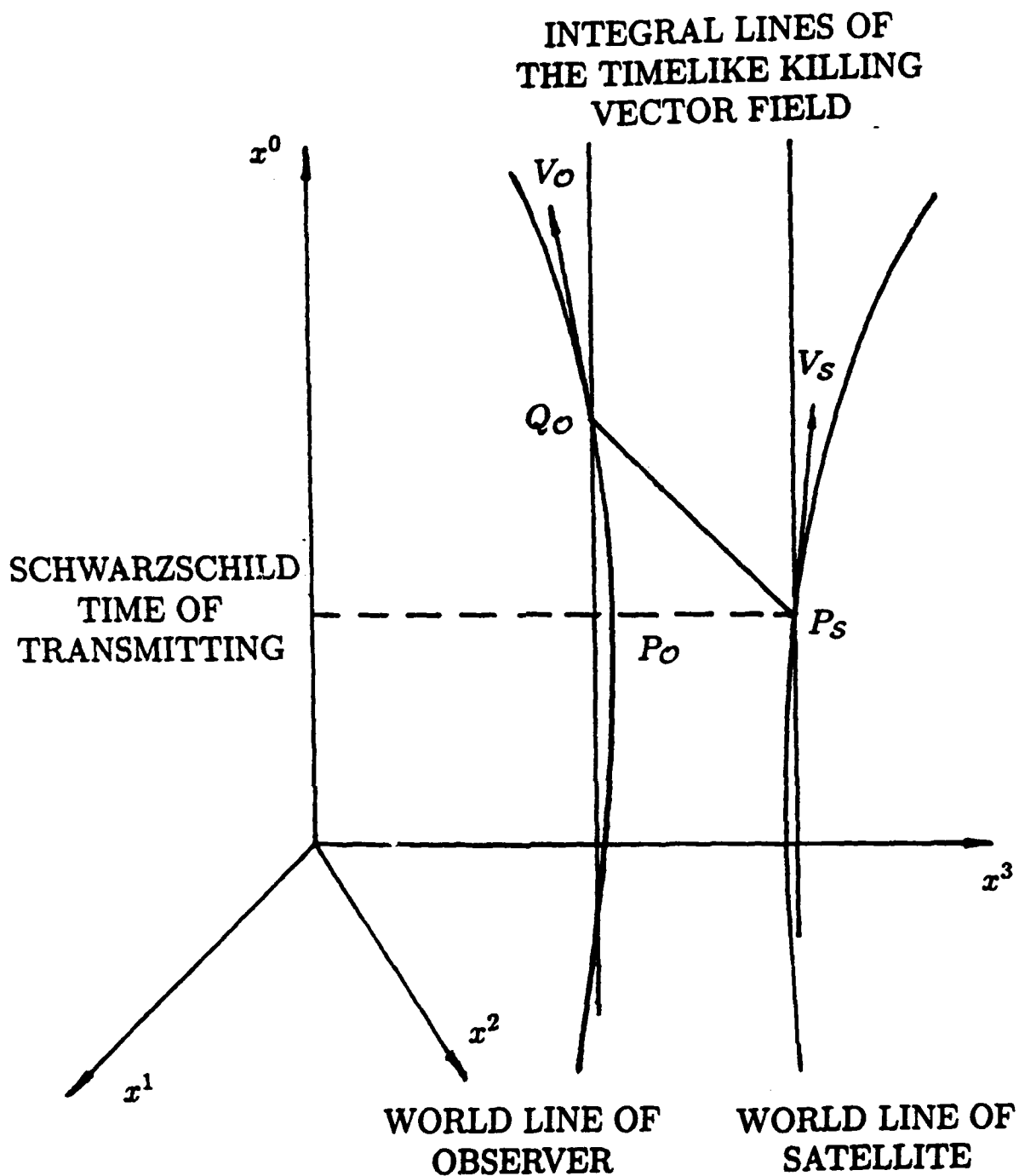
Fig. 2. The Doppler shift and its main contributing factors as viewed by observers resting with respect to the Schwarzschild coordinates.

The first two terms in this expression are the first order Doppler shift and the second order correction to the first order term (note that the second term is not symmetric with respect to V_O, V_S; it is related to the nonsymmetry of expression (2) with respect to ν_O, ν_S). The third term is the gravitational Doppler shift (cf. Eqn.. (4)). The last term can be called the centrifugal correction term, because, in a classical picture of circular orbits, the term can be thought of as the difference of potentials of centrifugal forces caused by the angular velocities of the satellite and the observer orbital motion. The physical origin of the last term in the general case is the motion of the satellite and the observer with respect to Schwarzschild coordinates.

We want to point out that in the case of circular orbits only the first two terms contain information about time delay between transmitting and receiving, and only these two terms are time dependent.

We also notice that Eqn. (5) coincides with the expression obtained by Ashby[4], excluding the gravitational term which has a general relativistic origin (Ashby's calculations were special relativistic).

IV. GLOBAL CLOCK SYNCHRONIZATION IN SCHWARZSCHILD FIELD

The relation between the clock rates and the Doppler shift is established[5,6] via the relation (cf. Fig. 3)

$$\mathcal{D} = \frac{\nu_S - \nu_O}{\nu_S} = 1 - \frac{\nu_O}{\nu_S} = 1 - \frac{d\tau_S}{d\tau_O}, \tag{6}$$

or

$$d\tau_S = (1 - \mathcal{D})d\tau_O \tag{7}$$

However, a closer look at this formula and at Fig. 3 makes it obvious that the infinitesimal interval of the satellite and the observer proper times $(d\tau_S, d\tau_O)$ are measured at different Schwarzschild times. A more precise form of (7) would be

$$(d\tau_S)_{t_t} = (1 - \mathcal{D})(d\tau_O)_{t_r}, \tag{8}$$

where t_t and t_r re the Schwarzschild times of transmission and reception of the signal. The retardation of t_r compared to t_t is reflected in Eqn. (5) by the structure of the first order term and the second order correction to the first order term. This circumstance was obviously the prime concern of H. Fliegel[2] of Aerospace Corporation.

The procedure, described above, of comparing the proper time rates of two clocks in general relativity is the only one (up to equivalence) that is correct for arbitrary gravitational fields. However, generally speaking, it will work only for two clocks (and under some reasonable conditions). In a general gravitational field (with no symmetries) it will not provide global synchronization for more than two clocks. It is not a drawback of this particular procedure. It is well known that in general relativity global synchronization of clocks in gravitational field with no symmetries is impossible in principle.
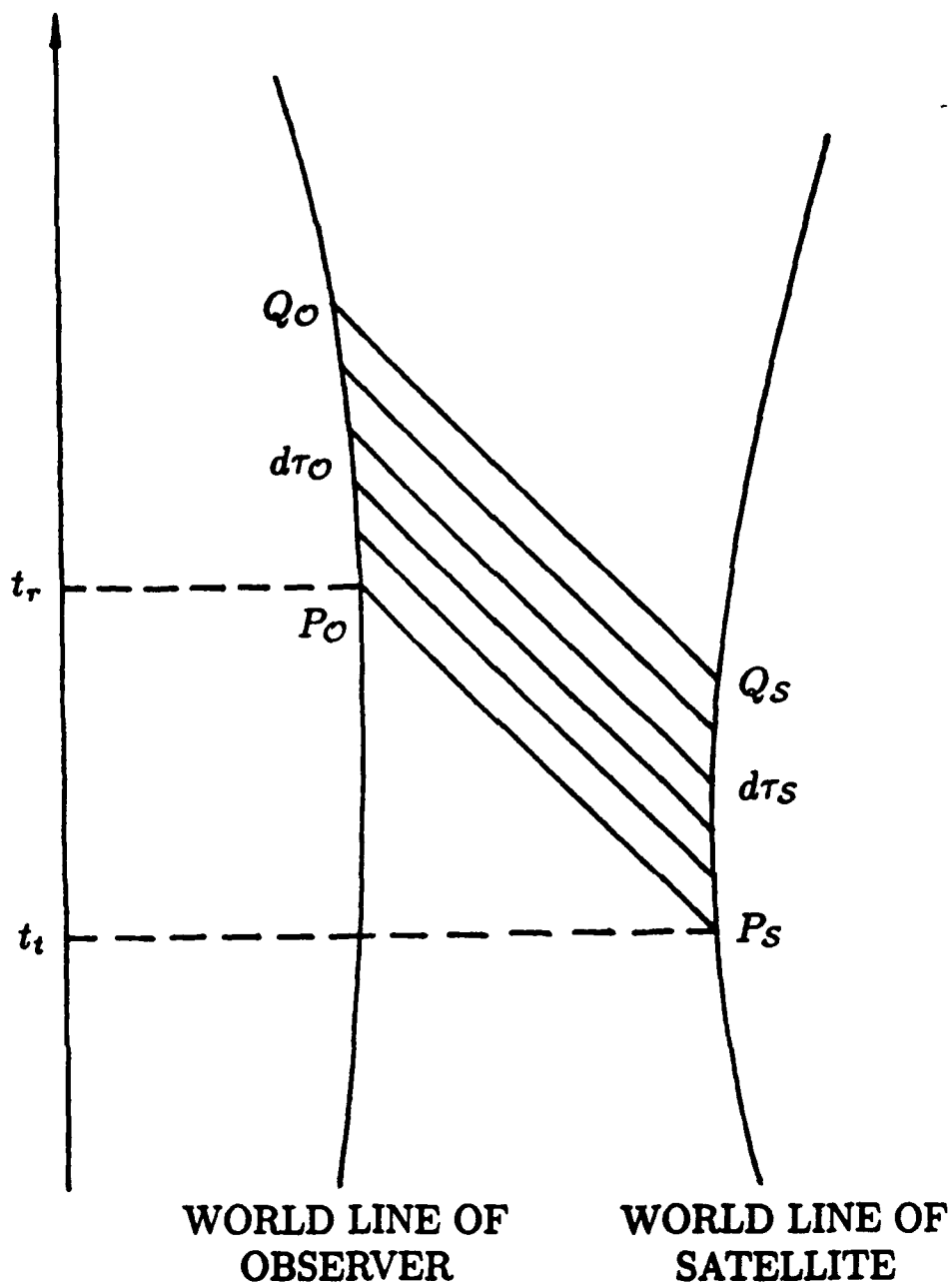
Fig. 3. Relation between the Doppler shift and the rates of the moving clocks. Shown is the set of null geodesics joining the world lines of the satellite and the observer. Each geodesic represents a wave crest. If there are n such crests and $d\tau_S$, $d\tau_O$ are the clock-measures of $P_S Q_S$ and $P_O Q_O$ respectively, then $n = \nu_S d\tau_S = \nu_O d\tau_O$

However, our model gravitational field of the earth (Schwarzschild field) is very symmetric (static, spherically symmetric). One can convince himself easily that in this particular case our procedure will do the job. But so will many others. The task is to find the simplest one. For instance, one would like to minimize the participation in the procedure of time dependent contributions like the first two terms of Eqn. (5). It would be a good idea to make all the clocks to display Schwarzschild coordinate time, i. e. the time of an observer placed at spatial infinity and resting with respect to the Schwarzschild coordinates. Schwarzschild coordinate time is the closest possible analog of the time of the ECI frame (the special relativistic limit of the Schwarzschild coordinate frame coincides with the ECI frame).

The first step in this direction is to compare the rates of the clocks of the satellite and the observer with Schwarzschild clocks simultaneously with respect to Schwarzschild time (cf. Fig. 4). Elementary calculations show that up to the second order
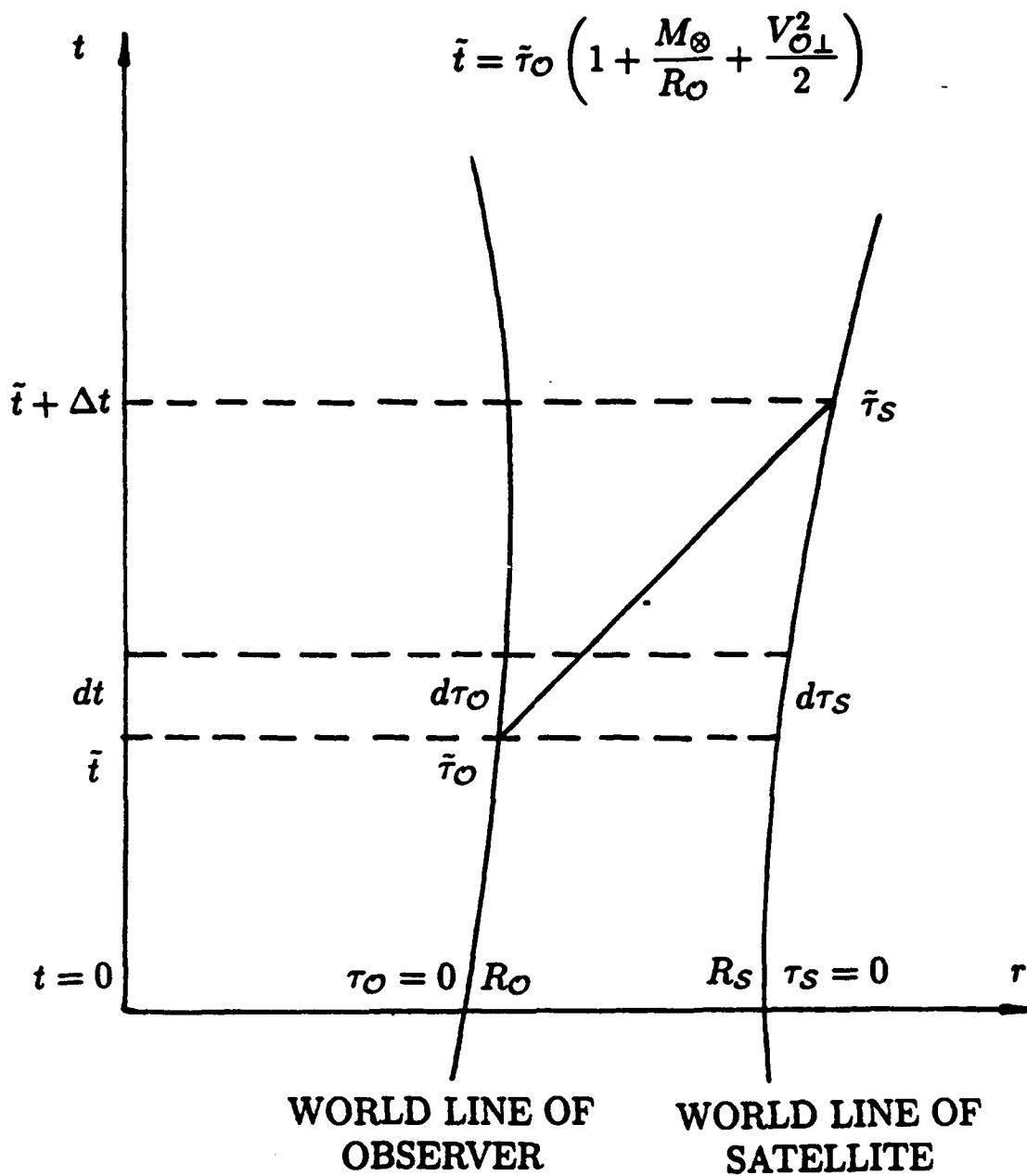
$$(d\tau_S)_t = \left(1 - \frac{2M_\otimes}{R_S}\right)^{\frac{1}{2}} (1 - V_{S\perp}^2)^{\frac{1}{2}} \, dt \qquad (9)$$

$$(d\tau_O)_t = \left(1 - \frac{2M_\otimes}{R_O}\right)^{\frac{1}{2}} (1 - V_{O\perp}^2)^{\frac{1}{2}} \, dt \qquad (10)$$

where $V_{O\perp}$ ($V_{S\perp}$) is the component of the observer's (satellite's) 4–velocity V_O (V_S) orthogonal to the timelike Killing vector field of the Schwarzschild metric.

We will perform the rest of our calculation in this section in a highly idealized fashion. Namely, we assume that R_O, $V_{O\perp}^2$, R_S, $V_{S\perp}^2$ are constant (the purpose of this idealization is to get rid of all the details of nonrelativistic origin). In this case both $d\tau_O$ and $d\tau_S$ are proportional to dt with constant proportionality coefficients. Thus one can take any of them as fundamental (the different choices are equivalent to the different choices of time units). Dividing (9) by (10), we obtain

$$(d\tau_S)_t = \frac{\left(1 - \frac{2M_\oplus}{R_s}\right)^{\frac{1}{2}} (1 - V_{S\perp}^2)^{\frac{1}{2}}}{\left(1 - \frac{2M_\oplus}{R_o}\right)^{\frac{1}{2}} (1 - V_{O\perp}^2)^{\frac{1}{2}}} (d\tau_O)_t, \qquad (11)$$

$$\tilde{t} = \tilde{\tau}_{\mathcal{O}} \left(1 + \frac{M_\otimes}{R_\mathcal{O}} + \frac{V_{\mathcal{O}\perp}^2}{2} \right)$$

Fig. 4. The relation between the satellite and the observer clock rates in a Schwarzschild simultaneity band and the initial clock synchronization procedure.

$$\Delta t = (R_S - R_\mathcal{O}) + 2M_\otimes \ln \frac{R_S - 2M_\otimes}{R_o o - 2M_\otimes}$$

or, in the usual second order approximation,

$$(d\tau_S)_t = \left[1 - \left(\frac{M_\oplus}{R_S} - \frac{M_\oplus}{R_O}\right) - \frac{1}{2}\left(V_{S\perp}^2 - V_{O\perp}^2\right)\right](d\tau_O)_t, \qquad (12)$$

which is interesting to compare with Eqn. (5) (note the loss of the terms related to the time delay).

Integration of Eqn. (12) yields

$$\tau_S = \left[1 - \left(\frac{M_\oplus}{R_S} - \frac{M_\oplus}{R_O}\right) - \frac{1}{2}\left(V_{S\perp}^2 - V_{O\perp}^2\right)\right]\tau_O + C. \qquad (13)$$

The constant of integration C can be made equal to zero by employing an appropriate choice of the origin for τ_O, τ_S. We will show one way to do it for the particular case when the observer is placed on equator, the plane of the satellite orbit is equatorial, and the orbit period is shorter than the period of the earth rotation. Let us suppose now that the ground observer is sending messages of his clock time continuously in the upward direction(we assume here that the aberration problem is properly taken care of), so that the satellite receiver knows that the signals are propagated along the radial null geodesics. For such signals expression (1) for the Schwarzschild metric implies (with $ds^2 = 0$, $d\theta = d\phi = 0$)

$$dt = \left(1 - \frac{M_\oplus}{r}\right)^{-1} dr. \qquad (14)$$

Integrating (14) we come up with the expression for the Schwarzschild travel time of the signal

$$\Delta t = R_S - R_O + 2M_\oplus \ln \frac{R_S - 2M_\oplus}{R_O - 2M_\oplus}. \qquad (15)$$

Thus, if the satellite receives the ground station message sent at $\tau_O = \tilde{\tau}_O$ and, at the moment of receiving, sets on its clock time to

$$\tilde{\tau}_S = \left(1 - \frac{M_\oplus}{R_S} - \frac{V_{S\perp}^2}{2}\right)\left[\left(1 + \frac{M_\oplus}{R_O} + \frac{V_{O\perp}^2}{2}\right)\tilde{\tau}_O + \Delta t\right], \qquad (16)$$

then the event on the world line of the satellite at $\tau_S = 0$ and on the world line of the observer at $\tau_O = 0$ become simultaneous with respect to Schwarzschild time,

and, if we choose as $t = 0$ the Schwarzschild time hypersurface passing through both events, then at any Schwarzschild moment of time t the clock of the observer and the satellite will display

$$\tau_O = \left(1 - \frac{M_\oplus}{R_O} - \frac{V_{O\perp}^2}{2}\right) t, \tag{17}$$

and

$$\tau_S = \left(1 - \frac{M_\oplus}{R_S} - \frac{V_{S\perp}^2}{2}\right) t, \tag{18}$$

making it possible to tell Schwarzschild time by looking at any of the clocks.

The constant C in Eqn. (13) thus becomes equal to zero. It is clear that the described procedure allows one to synchronize as many clocks as he wishes to Schwarzschild time and the procedure can be generalized to any placement of the observer on the earth and any satellite orbit inclination. In so doing, one might expect the expression for Δt to become more complicated.

Therefore, it is worth to take another look of the Eqn. (15) and to evaluate the last term in it since this term is the one that has a tendency to become more complicated. Let us rewrite Eqn. (15) as follows

$$\Delta t = R_S - R_O + KR_O \tag{19}$$

where

$$K = \frac{2M_\oplus}{R_O} \left(\ln \frac{R_S}{R_O} + \ln \frac{1 - 2M_\oplus/R_S}{1 - 2M_\oplus/R_O}\right) \tag{20}$$

Using the Maclaurin series expansion and dropping all the powers of $M_\oplus/r$ higher than the first, we estimate (20) as follows

$$K \approx \frac{2M_\oplus}{R_O} \left(\ln \frac{R_S}{R_O} - \frac{2M_\oplus}{R_S} + \frac{2M_\oplus}{R_O}\right) \approx \frac{2M_\oplus}{R_O} \ln \frac{R_S}{R_O} \tag{21}$$

Taking into account that for the GPS satellites $R_S \approx 2R_O$, and using the values of $M_\oplus$ and R_O, we come up with $K \sim 10^{-9}$. This means that if we rep.ace the exact expression (19) for an approximate one

$$\Delta t \approx R_S - R_O \tag{22}$$

we introduce an error in ranging of the order of 1 cm. The error is introduced in the constant of integration and, consequently, does not accumulate. It is clear, therefore, that for any practical purpose we can neglect the last term and use the approximate range–time–data relation (22) instead of exact equation (19).

It is interesting to note here that Eqn. (22) can be also written as

$$\Delta t = \Delta r (1 + O_2) \tag{23}$$

This gives us a hint that the range–data relation might admit the general relativistic interpretation and that it is satisfied only up to the first order. This subject will be developed in a general context and in more detail in section VI.

However, we can make a conjecture that, most probably, the initial synchronization procedure can use the special relativistic range–time data relation in present-day GPS. The conjecture, as any conjecture, is formulated in a rather vague fashion. We will turn it into a precise statement in subsequent sections.

V. GEOMETRIC APPROACH AND THE "RELATIVITY OF SIMULTANEITY".

As it is well known[8,9] the geometric approach is a key feature of modern relativity and its applications. This approach helps one to see clearly the physics of relativistic procedures at all steps of calculations and, typically, it considerably reduces the amount of calculations necessary to describe relativistic effects. We are going to demonstrate the power of the approach by analyzing the proposal to change the relativistic treatment of the GPS time transfer by means of the "taking into account" of the so-called "relativity of simultaneity"[6] (we will use the abbreviated expression "RS-proposal" for it in the rest of the discussion). This proposal led to an incredible waste of time and effort of the researchers in 1987-1988. It was based heavily on the special relativistic concept of simultaneity with respect to an inertial frame of reference applied mistakenly to non-inertial frames. This concept has no analogues in general relativity. The techniques of calculations used in this approach was that of the pre-Minkowski epoch thus leading to many pages of calculations and creating ample opportunity for producing mistakes.

The geometry of the RS-proposal, as we show below, is primitive. If the geometric approach had been used, the main results of the RS-proposal[6] could have been obtained in two lines (cf. Eqns. (29)-(30) below), together with a clear physical interpretation of what actually had been done.

To compare the general relativistic calculations of section IV with the special relativistic calculations of the RS-proposal we consider the case $M_{\oplus} = 0$ in the equations of section IV. This means that we neglect the gravitational correction. Thus the Schwarzschild frame transforms into the ECI frame. Therefore,

$$\Delta t = R_S - R_O, \tag{24}$$

$$d\tau_O = (1 - \vec{V}_O^2)^{\frac{1}{2}}, \tag{25}$$

and

$$dτ_S = (1 - \vec{V}_S^2)^{\frac{1}{2}}. \tag{26}$$

(we have replaced $V_\perp$ of section IV for $\vec{V}$; $V_\perp$ and $\vec{V}$ coincide up to the second order). Let us try to reconstruct now the geometry of the RS-proposal. It can be expressed as follows.

The RS-proposal suggests that we compare the intervals of proper times $dτ_S$ and $dτ_O$ simultaneous in the frame of the observer (cf. Fig. 5),

$$dτ_O = -(V_O \cdot V_S)dτ_S, \tag{27}$$

where the components of the 4-velocities V_O and V_S, as represented in the ECI frame basis, are given by (up to the second order),

$$V_O = \left((1 - \vec{V}_O^2)^{-\frac{1}{2}}, \vec{V}_O\right), \qquad V_S = \left((1 - \vec{V}_S^2)^{-\frac{1}{2}}, \vec{V}_S\right). \tag{28}$$

Here $\vec{V}_O$ and $\vec{V}_S$ are the 3-velocities of the observer and the satellite, respectively, in the ECI frame. Substituting (28) into (27) we obtain (up to the second order),

$$(dτ_O)_{t_2} = \left((1 - \vec{V}_O^2)^{-\frac{1}{2}}(1 - \vec{V}_S^2)^{-\frac{1}{2}} - \vec{V}_O \cdot \vec{V}_S\right)(dτ_S)_{t_1}, \tag{29}$$

or, otherwise,

$$(dτ_S)_{t_1} = \left(1 - \frac{1}{2}(\vec{V}_O^2 + \vec{V}_S^2) + \vec{V}_O \cdot \vec{V}_S\right)(dτ_O)_{t_2}. \tag{30}$$

It is instructive to look now at the geometry of the RS-proposal. What happens is that we are using for the ordering in time the alleged slicing of spacetime by the family of proper spaces of the observer, i. e. by planes $τ_O = $ const, instead of the slicing by proper spaces of ECI frame. But the family of the proper spaces of the observer does not provide a slicing (cf. Fig. 6). These proper spaces intersect each other. As a result, the global synchronization has not been achieved (and cannot be achieved) in this way. Also, one should notice that $\vec{V}_O$ and $\vec{V}_S$ are not measured
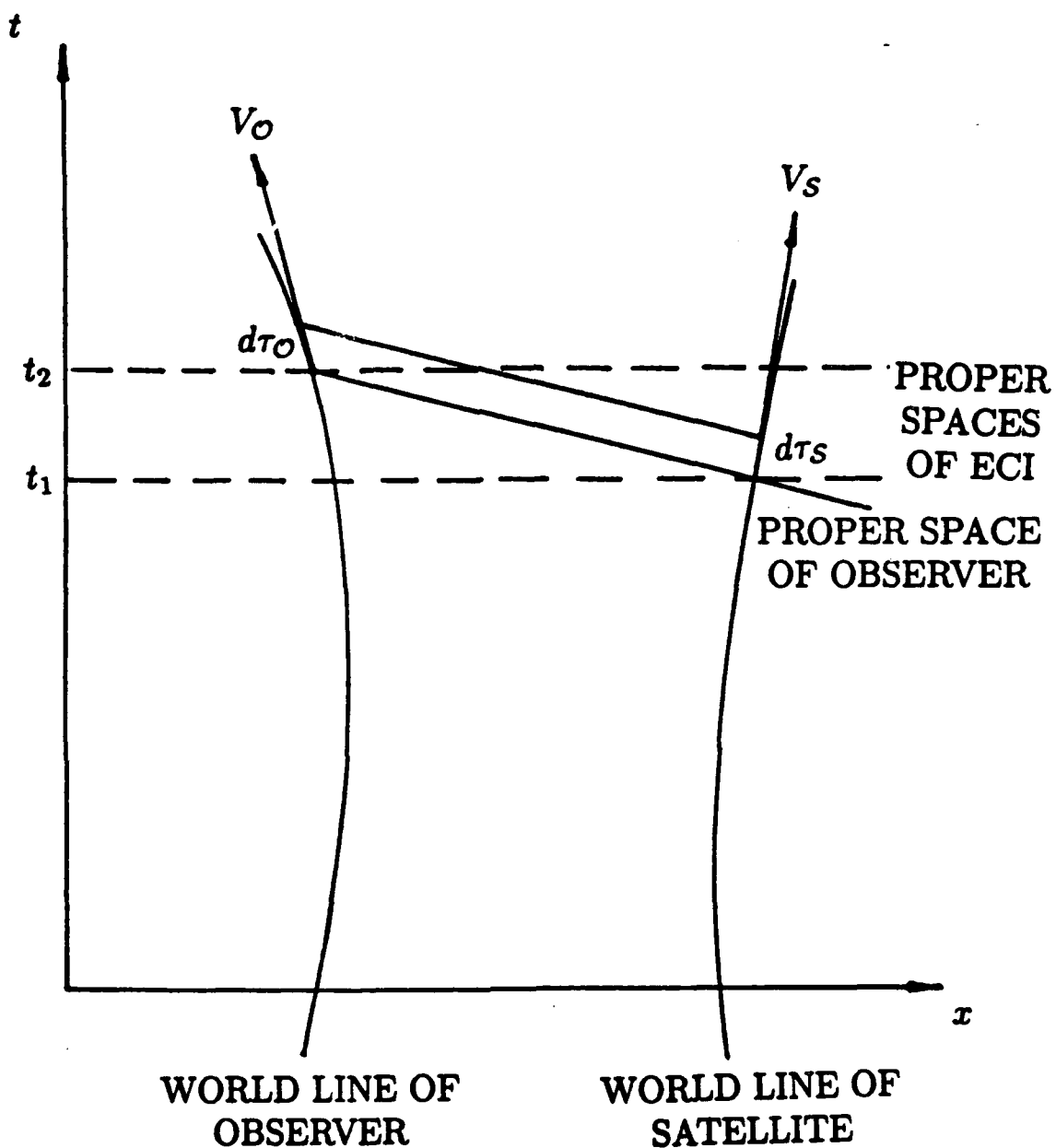
Fig. 5. The geometry of the "relativity of simultaneity" proposal. The intervals of the satellite and the observer proper time are compared not in the Schwarzschild simultaneity band but rather in the simultaneity band of the instantaneous comoving frame of the observer. Such a construction cannot be defined in curved spacetime.
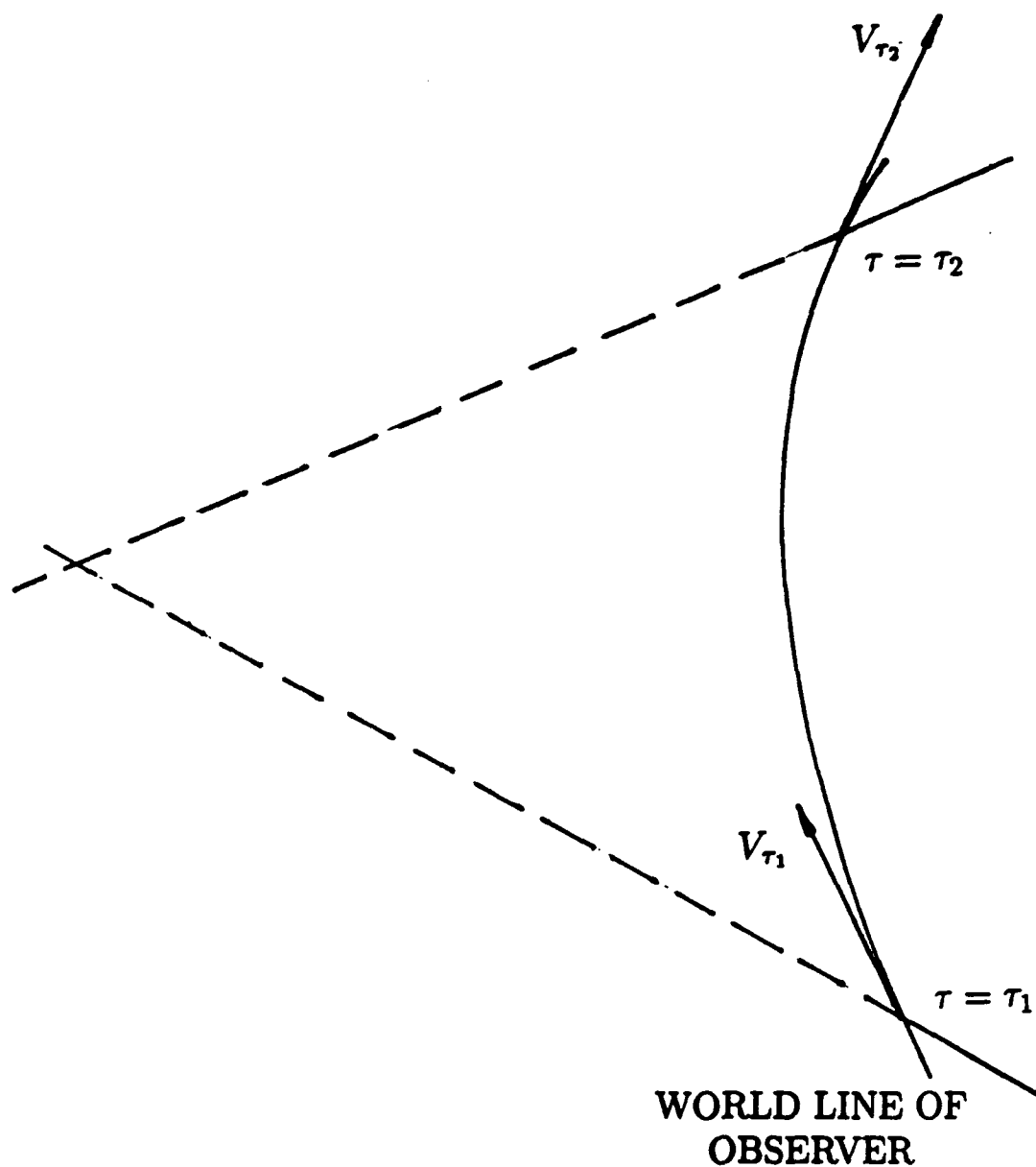
Fig. 6. Even in special relativity proper spaces of the instantaneous comoving frame of the observer are not capable of supporting global synchronization. Acceleration of the observer causes the proper spaces of the instantaneous comoving frames of the observer corresponding to two different moments of time to intersect each other. The simultaneity bands are not well defined in this approach.

at the same moment of Schwarzschild time, so that the scalar product in equations (29) and (30) actually should be written as $(\vec{V}_O)_{t_2} \cdot (\vec{V}_S)_{t_1}$ (one should perform similar corrections in other terms; however, in the case of circular orbits $\vec{V}_O^2$ and $\vec{V}_S^2$ are constant).

Expressions (29) and (30) are not useful for comparing proper times τ_O and τ_S with the ECI frame time t. Even if we decided to do so (in a close neighborhood of the observer world line), it would require taking into account (consisting of bulky calculations) the difference between t_1 and t_2 at each moment when integrating, and not just for the initial synchronization.

The equations themselves are correct. Nevertheless, they have nothing to do with synchronization. One can use them (close to the observer world line) if one finds out the relation between dt_1 and dt_2, which is not hard to do. The result is,

$$dt_2 = \left[(1 - \vec{V}_O^2)^{-\frac{1}{2}} (1 - \vec{V}_S^2)^{-\frac{1}{2}} - \vec{V}_O \cdot \vec{V}_S \right]^{-1} \frac{(1 - \vec{V}_S^2)^{\frac{1}{2}}}{(1 - \vec{V}_O^2)^{\frac{1}{2}}} dt_1, \qquad (31)$$

which, when used together with Eqns. (29)–(30), gives

$$(d\tau_S)_{t_1} = \frac{(1 - \vec{V}_S^2)^{\frac{1}{2}}}{(1 - \vec{V}_O^2)^{\frac{1}{2}}} (d\tau_O)_{t_1} = \left(1 - \frac{1}{2}(\vec{V}_S^2 - \vec{V}_O^2) \right) (d\tau_O)_{t_1}. \qquad (32)$$

This is identical to the analysis presented in section IV. The difficulties of the RS-proposal are related mainly to the missing piece of information, namely (cf. Fig. 7),

$$dt_1 \neq dt_2 \neq dt_3. \qquad (33)$$

Otherwise, this analysis would work as well as any other correct procedure, although it is not clear why one should put himself through all of this to get simple results. Of course, it wipes out the effect of the cross-term (the key result of the RS-approach).

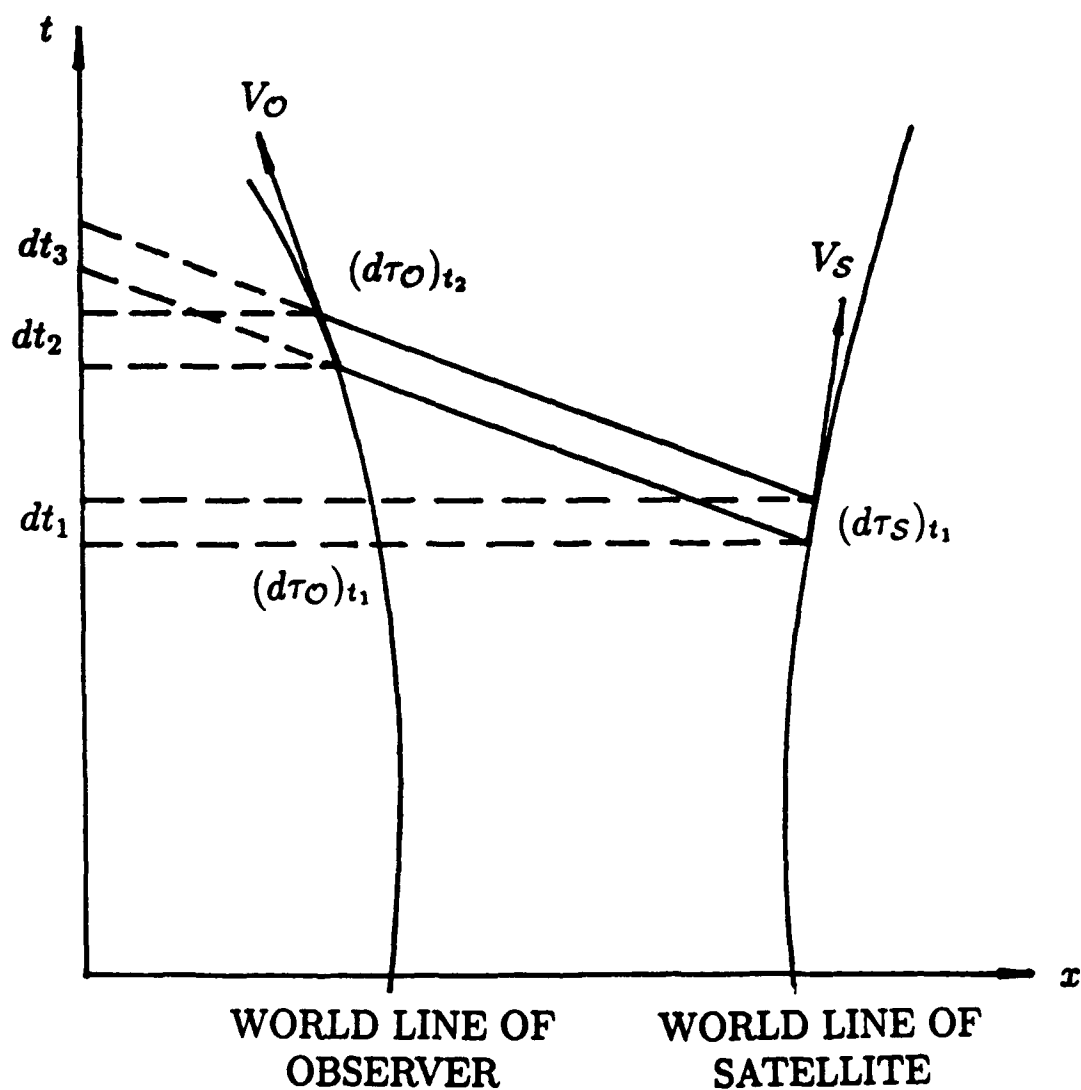Fig. 7. The time intervals dt_1, dt_2, and dt_3 are not equal. After taking into account the correct relation between them, we come up with the clock synchronization scheme equivalent to the synchronization in the Schwarzschild coordinate system.

VI. RELATION BETWEEN THE RANGE DATA AND THE TIME DATA IN A SCHWARZSCHILD FIELD.

We are going to consider now the relation between the range and time data in a Schwarzschild field for the light (or radio) signal traveling between the ground observer and the satellite. The placement of the ground observer is not restricted anymore. He can be placed anywhere at the ground surface and may even be in motion. The satellite orbital plane can be inclined with respect to the equatorial plane and the orbit does not have to be circular. Actually, the analysis that we are going to undertake can be applied to a pair of satellites or even to a pair of spaceships which are accelerating while the ranging is being performed.

Mathematically the problem can be formulated as follows. First we introduce the coordinates (x^0, x^1, x^2, x^3)

$$x^0 = t, \quad x^1 = r \sin \theta \cos \phi, \quad x^2 = r \sin \theta \sin \phi, \quad x^3 = r \cos \theta \tag{34}$$

the same way as we did in section III.

The Schwarzschild metric reexpressed in these coordinates is

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu = - \left(1 - \frac{2M_\oplus}{r} \right) dx^{0^2} + dx^k dx^k$$
$$+ \frac{2M_\oplus}{r^3} \left(1 - \frac{2M_\oplus}{r} \right)^{-1} (x^k dx^k)^2 \tag{35}$$

where $k = 1, 2, 3$, and the summation over repeating indices is assumed. It is interesting to note here that the Schwarzschild radial coordinate r can be expressed as

$$r^2 = x^k x^k \tag{36}$$

Up to second order we have

$$g_{\mu\nu} = \eta_{\mu\nu} + \gamma_{\mu\nu} \tag{37}$$

$$\eta_{\mu\nu} = \text{diag}(-1, 1, 1, 1) \tag{38}$$

$$\gamma_{00} = \frac{2M_\oplus}{r} \qquad (39)$$

$$\gamma_{0k} = 0 \qquad (40)$$

$$\gamma_{ik} = \frac{2M_\oplus x^i x^k}{r^3} = 2M_\oplus(r^{-1}\delta_{ik} - r_{,ik}) \qquad (41)$$

where δ_{ik} is the Kronecker delta and $r_{,ik} = \frac{\partial^2 r}{\partial x^i \partial x^k}$.

We want to stress that in the space of these coordinates (pictured as if they were Cartesian coordinates) the geodesic lines and, particularly, the null geodesic lines do not usually look like straight lines.

Let us suppose (cf. Fig. 8) that in this space $\mathcal{O}$ is the event of emitting the light (or radio) signal by the ground observer, and $\mathcal{S}$ is the event of receiving the signal by the satellite. The solid (curved) line $\mathcal{OS}$ represents the null geodesic world line of the photon traveling from $\mathcal{O}$ to $\mathcal{S}$. The vertical dashed lines represent the integral lines of the timelike Killing vector field of the Schwarzschild metric passing through the events $\mathcal{O}$ and $\mathcal{S}$. In general

$$\Delta t \neq (\Delta x^k \Delta x^k)^{\frac{1}{2}} \qquad (42)$$

If there was no gravitational field, i. e. if the spacetime was flat the world line of the photon received at $\mathcal{S}$ would be the straight inclined dashed line $\mathcal{O}'\mathcal{S}$ intersecting the integral line of the timelike Killing vector field passing through $\mathcal{O}$ at a point $\mathcal{O}'$ not coinciding with $\mathcal{O}$. Then we would have

$$(\Delta t)_f = (\Delta x^k \Delta x^k)^{\frac{1}{2}} \qquad (43)$$

The straight line segment $\mathcal{OO}'$ pictures the difference between Δt and $(\Delta t)_f$, and it is this difference

$$\Delta t - (\Delta t)_f = \Delta t - (\Delta x^k \Delta x^k)^{\frac{1}{2}} \qquad (44)$$
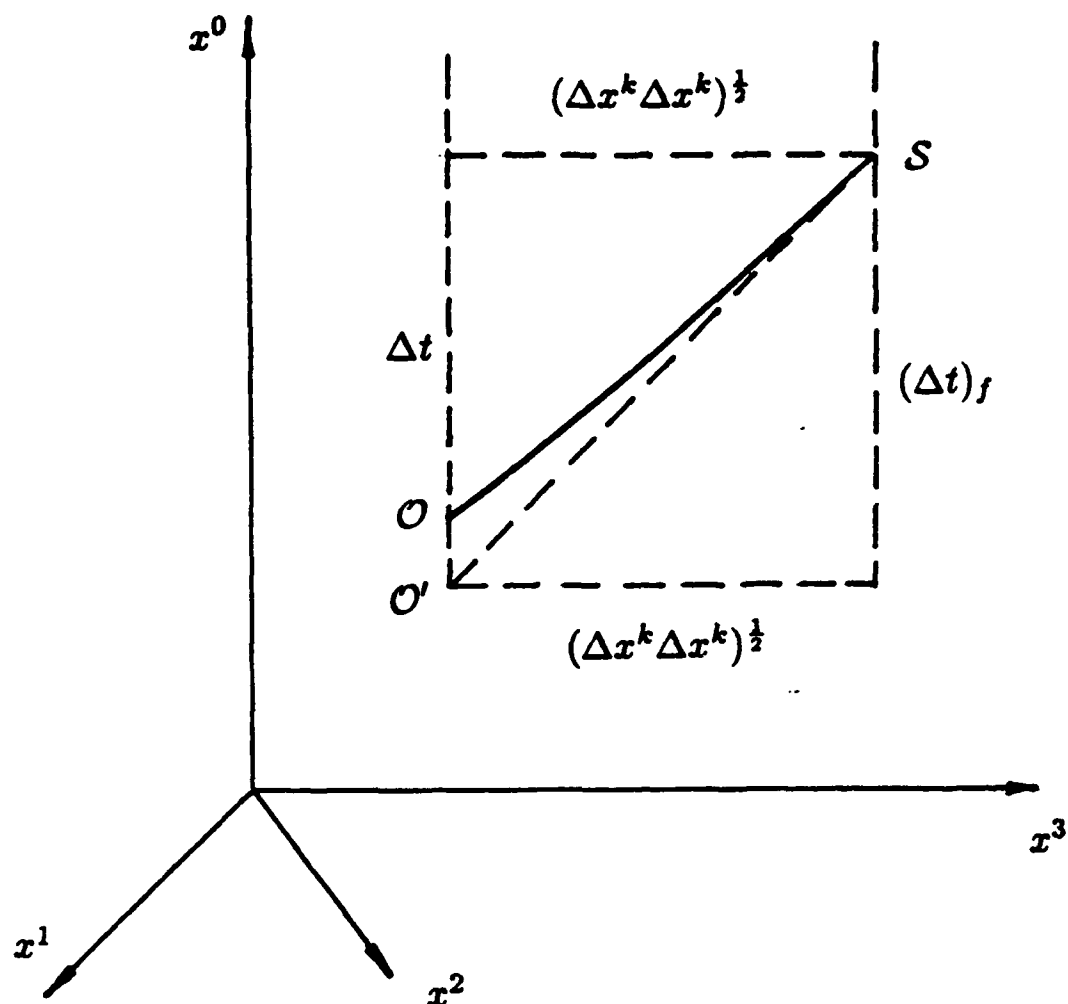
Fig. 8. An estimate of the difference between the general relativistic and special relativistic range–time relation. The special relativistic relation can be used, but it introduces an error of the second order with respect to $\sqrt{M_\oplus/r}$.

that we want to estimate.

Let us recall that the line OS is a null geodesic. Its equations can be written as

$$\frac{d^2x^\mu}{dw^2} + \Gamma^\mu_{\lambda\nu}\frac{dx^\lambda}{dw}\frac{dx^\nu}{dw} = 0 \tag{45}$$

where $\Gamma^\mu_{\lambda\nu}$ are Kristoffel symbols and w is an affine parameter such that $w = 0$ at O and $w = 1$ at S. Calculations using the covariant tensor series expansion (J. L. Synge's world function[9]) lead us to the equation (precise up to the second order)

$$\eta_{\mu\nu}\Delta x^\mu \Delta x^\nu = Q \tag{46}$$

where

$$Q = s\gamma_{\mu\nu}\Delta x^\mu \Delta x^\nu - 2\Delta x^\mu \Delta x^\nu \int_0^1 \gamma_{\mu\nu}dw - \Delta x^\mu \Delta x^\nu \Delta x^\lambda \int_0^1 \gamma_{\mu\nu,\lambda}wdw \tag{47}$$

Hence

$$\Delta t^2 = \Delta x^k \Delta x^k - Q \tag{48}$$

$$\Delta t = (\Delta x^k \Delta x^k)^{\frac{1}{2}} - \frac{1}{2}Q(\Delta x^k \Delta x^k)^{-\frac{1}{2}} \tag{49}$$

The last term in (49) is small, so that in calculation of Q we may substitute $\Delta t = (\Delta x^k \Delta x^k)^{\frac{1}{2}}$. Eqn. (49) can be rewritten in form

$$\Delta t = (\Delta x^k \Delta x^k)^{\frac{1}{2}}\left[1 - \frac{1}{2}Q(\Delta x^k \Delta x^k)^{-1}\right] \tag{50}$$

or, introducing notation

$$F = -\frac{1}{2}Q(\Delta x^k \Delta x^k)^{-1} \tag{51}$$

reduced to

$$\Delta t = (\Delta x^k \Delta x^k)^{\frac{1}{2}}[1 + F] \tag{52}$$

The expression (51) can be evaluated but it is not an easy task. However the order of magnitude of F can be estimated rather easily for the case $\Delta x^k \Delta x^k \sim r^2$ which

is the case in all GPS problems. Just by looking at Eqns. (52), (47), and (39)–(41) one can make an obvious conclusion

$$F \sim \frac{M_\oplus}{r} \qquad (53)$$

This solves clearly the objectives (2) and (3) of section II. The standard GPS range – time – data relation takes form

$$\Delta t = (\Delta x^k \Delta x^k)^{\frac{1}{2}} \qquad (54)$$

The relation indeed can be interpreted globally in the region of curved Schwarzschild spacetime where $M_\oplus/r$ is small, provided that t is Schwarzschild time and the coordinates x^k are related to the Schwarzschild coordinates as in Eqns. (34). The relation (54) is approximate. It is satisfied only up to the first order (with respect to $\sqrt{M_\oplus/r}$).

This result does not depend in any way on 4–velocities and accelerations of the ground observer and the satellite, so that it remains true also for a pair of satellites or even for a pair of spaceships with arbitrary accelerations.

It is also clear that the procedure of initial synchronization described at the end of section IV will work in the general case of the observer placement and an arbitrary choice of the satellite orbit. All that one needs to do is to relax the requirement that the signal should be sent vertically (one can send it any way he wants), and to use our new relation (54) instead of the range – time –data relation (22). In this way, he will be able to perform the initial clock synchronization (setting the constant of integration in the Eqn. (13) to zero) with an error of the first order. Since the error does not accumulate in time this synchronization is quite sufficient for any practical purpose of modern GPS.

VII. NONCIRCULAR ORBITS IN GENERAL RELATIVITY.

In section IV we obtained the expression relating the time of the ground observer clock and the satellite clock for the case of the satellite circular orbit (Eqn. (13)). One can notice, however, that all the analysis preceding the Eqn. (13) does not use the assumption of a circular orbit. Eqn. (12) remains unchanged if we relax this requirement. It is the procedure of integration of the Eqn. (12) leading to the Eqn. (13) that uses the assumptions R_S = const and $V_{S\perp}^2$ = const (which are equivalent to assuming the satellite orbit being circular).

Although the orbits of the GPS satellites are meant to be circular in reality they can never be perfectly circular. We want to know now what happens if the satellite orbit is slightly noncircular.

In Newtonian mechanics the orbits slightly deviating from circular are elliptic. The simplest way to analyze the satellite motion on such an elliptic orbit is to use the Hamilton–Jacobi method[8]. The main results are as follows

(1) The orbit of the satellite is planar. Consequently, one can introduce spherical coordinates (r, θ, ϕ) in such a way that

$$\theta = \frac{\pi}{2} = \text{const} \tag{55}$$

(2) The total energy of the satellite and its angular momentum are conserved and so are the total energy per unit mass of the satellite $\tilde{\varepsilon}$ and its angular momentum per unit mass $\tilde{L}$

$$\tilde{\varepsilon} = -\frac{M_\oplus}{r} + \frac{V^2}{2} \tag{56}$$

$$\tilde{L} = r^2 \dot{\phi} \tag{57}$$

(3) The orbit of the satellite is elliptic and it is defined by the equation

$$\theta = \int \frac{\tilde{L} \left[2(\tilde{\varepsilon} + M_\oplus/r - \tilde{L}^2/2r^2) \right]^{-\frac{1}{2}} dr}{r^2} \tag{58}$$

or, after integration,

$$r = \frac{\tilde{L}^2/M_\oplus}{1 + e \cos \phi} \qquad (59)$$

where e is the eccentricity of the orbit

$$e = \left(1 + \frac{2\tilde{\varepsilon}\tilde{L}^2}{M_\oplus^2}\right)^{\frac{1}{2}} \qquad (60)$$

The semimajor axis of the orbit a is

$$a = r_{\text{max}} - r_{\text{min}} = \frac{\tilde{L}^2/M_\oplus}{1 - e^2} = \frac{M_\oplus}{(-2\tilde{\varepsilon})} \qquad (61)$$

The constant of integration in (58) has been picked up in such a way that the position of the closest approach (periastron) is achieved when $\phi = 0$. The satellite returns to the periastron position at $\phi = 2k\pi$ for any integer k.

(4) Time as correlated with position is given by

$$t = \int \left[2(\tilde{\varepsilon} + M_\oplus/r - \tilde{L}^2/2r^2)\right]^{-\frac{1}{2}} dr \qquad (62)$$

To simplify the integration it is common practice to introduce a new parameter u so that

$$r = \frac{M_\oplus}{(-2\tilde{\varepsilon})^{3/2}}(1 - e \cos u) = a(1 - \cos u) \qquad (63)$$

The parameter u is the so-called "mean eccentric anomaly", or, otherwise, Bessel's time parameter. Substitution of (63) into (62) and subsequent integration gives

$$t = \frac{M_\oplus}{(-2\tilde{\varepsilon})^{3/2}}(u - e \sin u) \qquad (64)$$

where the constant of integration is chosen so that at $t = 0$, $u = 0$.

Bessel's time parameter is related to the angle coordinate ϕ as follows

$$\sin u = \frac{(1 - e^2)^{\frac{1}{2}} \sin \phi}{1 + e \cos \phi} \qquad (65)$$

$$\cos u = \frac{\cos \phi + e}{1 + e \cos \phi} \tag{66}$$

$$\cos \phi = \frac{\cos u - e}{1 - e \cos u} \tag{67}$$

$$\sin \phi = \frac{(1 - e^2)^{\frac{1}{2}} \sin u}{1 - e \cos u} \tag{68}$$

Although u has the same period as ϕ, the difference between them is very essential, apart from the case when $e = 0$ (circular orbit) which gives $u = \phi$. We will be interested later in the relation between u and ϕ for $e \neq 1$ when ϕ and u are small. Then (64) yields (up to the first order)

$$u \approx \frac{(1 - e^2)^{\frac{1}{2}} \phi}{1 + e} = \sqrt{\frac{1 - e}{1 + e}} \phi \tag{69}$$

which means that for small ϕ the parameter u is of the same or higher order of smallness as ϕ

$$u \sim \phi \tag{70}$$

In general relativity the satellite motion for a noncircular orbit is more complicated. In general, the motion is nonperiodic. The physical reason for the difference stems from the fact that, in general relativity, the period of the radial motion does not coincide with the period of the angular motion of the satellite. For the orbit of small eccentricity the difference between the classical an relativistic motion can be pictured as the periastron shift.

For a nearly circular orbit of the radius r_0, the angle swept between two successive periastrons is

$$\Delta \phi = 2\pi \left(1 - \frac{6 M_\otimes}{r_0} \right)^{-\frac{1}{2}} \tag{71}$$

or, up to the first order,

$$\Delta \phi \approx 2\pi + 6\pi \frac{M_\otimes}{r_0} \tag{72}$$

i.e., the angle variable acquires an additional shift

$$\delta\phi = 6\pi\frac{M_\oplus}{r_0} \tag{73}$$

per one period of ra.. .l motion compared to the classic case (when $\delta\phi = 0$).

As in classical mechanics a complete description of the satellite motion can be produced using the Hamilton–Jacobi method. The main results can be presented in the following way.

(1) The orbit of the satellite is planar, as in the case of classical mechanics, so that the Schwarzschild coordinates (t, r, θ, ϕ) can be chose in such a way that

$$\theta = \text{const} = \frac{\pi}{2} \tag{74}$$

(2) The total energy per unit mass $\tilde{E}$ and the angular momentum per unit mass $\tilde{L}$ are conserved. $\tilde{E}$ and and $\tilde{L}$ are called the energy (per unit mass) and the angular momentum (per unit mass) at infinity (for a Schwarzschild observer at infinity). We want to note that the Eqn. (57) is not satisfied exactly in general relativity but is still correct up to the second order.

(3) The orbit of the sa..ellite is defined by the equation

$$\phi = \int \frac{\tilde{L}\left[\tilde{E}^2 - (1 - 2M_\oplus/r)(1 + \tilde{L}^2/r^2)\right]^{-\frac{1}{2}} dr}{r^2} \tag{75}$$

(4) Time, as correlated with the position of the satellite, is given by

$$t = \int \tilde{E}\left[\tilde{E}^2 - (1 - 2M_\oplus/r)(1 + \tilde{L}/r^2)\right]^{-\frac{1}{2}} \frac{dr}{(1 - 2M_\oplus/r)} \tag{76}$$

It is obvious that to perform calculations using relations (75), (76) is much harder than for the classical formulae (58), (62). This circumstance motivates us to estimate first the difference in predictions from (75), (76) compared to those from (58), (62).

The detailed analysis shows that the difference between them is of second order. One of the ways to handle the situation is to modify the parameters. We will consider the details only for the particular case that we will use in this report, i. e. only for the Eqns. (62), (63). To keep these equation unchanged we modify the argument u for $u + \delta u$ where $\delta u \sim \frac{M_{\circledast}}{r}$, i. e.

$$r = a[1 - e\cos(u + \delta u)] \tag{77}$$

$$\frac{M_{\circledast}^{1/2}}{a^{3/2}}t = u + \delta u - e\sin(u + \delta u) \tag{78}$$

These equations coincide with the classical equations only up to the first order.

VIII. THE RATES OF CLOCKS FOR NONCIRCULAR ORBITS.

As we mentioned above the differential relations (9)–(12) of section IV are correct for an arbitrary orbit. It is only the integrated relation (13) that uses the assumption of the orbit being circular. Using the results of section VII we are going to modify the results of section IV to include the case of a slightly noncircular orbit. Let us rewrite Eqn. (9) in the form

$$(d\tau)_t = \left[1 - \frac{M_\otimes}{r} - \frac{V^2}{2}\right] dt \tag{79}$$

We have dropped the index S everywhere because we are discussing now only the satellite motion. We will restore it whenever it becomes necessary.

Let us modify now Eqn. (79) to the form more suitable for integrating in the case of noncircular orbits. First, we use the fact that the Newtonian equation of the total energy conservation per unit mass is also true in general relativity up to the second order

$$-\frac{M_\otimes}{r} + \frac{V^2}{2} = \tilde{\varepsilon} = -\frac{M_\otimes}{2a} = \text{const} \tag{80}$$

This allows us to eliminate $V^2/2$ from Eqn. (79)

$$(d\tau)_t = \left[1 - 2M_\otimes \left(\frac{1}{r} - \frac{1}{4a}\right)\right] dt \tag{81}$$

Now we can use Eqn. (77), from which it follows that up to the first order

$$\frac{1}{r} = \frac{1}{a}\frac{1}{1 - e\cos u} \tag{82}$$

Substitution of (82) in (81) produces a (correct up to second order) expression

$$(d\tau)_t = \left[1 - \frac{3}{2}\frac{M_\otimes}{a} - \frac{2M_\otimes e}{a}\frac{\cos u}{(1 - e\cos u)}\right] dt \tag{83}$$

where the Bessel's time parameter u is a function of t. Integrating Eqn. (83) we come up with

$$\tau = t - \frac{3}{2}\frac{M_\otimes}{a}t - \frac{2M_\otimes e}{a} \int \frac{\cos u}{1 - e\cos u} dt + C \tag{84}$$

To evaluate the term

$$\frac{2M_\otimes e}{a} \int \frac{\cos u}{1 - e\cos u} dt \qquad (85)$$

we use the fact that differentiation of Eqn. (78) yields, up to the first order

$$dt = \frac{a^{3/2}}{M_\otimes^{1/2}}(1 - e\cos u)du \qquad (86)$$

Thus, up to second order

$$\frac{2M_\otimes e}{a} \int \frac{\cos u}{1 - e\cos u} dt = 2M_\otimes^{\frac{1}{2}} a^{\frac{1}{2}} e \int \cos u\, du = 2M_\otimes^{\frac{1}{2}} a^{\frac{1}{2}} e(\sin u - \sin u_0) \qquad (87)$$

so that, finally, up to the second order, we have

$$\tau_S = t - \frac{3}{2}\frac{M_\otimes}{a}t - 2M_\otimes^{\frac{1}{2}} a^{\frac{1}{2}} e(\sin u - \sin u_0) + C_S \qquad (88)$$

For the ground observer the expression for τ_O does not change

$$\tau_O = t - \frac{M_\otimes}{R_O}t - \frac{V_{O\perp}^2}{2}t + C_O \qquad (89)$$

The initial synchronization procedure described at the end of section IV should be modified as follows. First, we define the global Schwarzschild coordinate time t so that $t = 0$ when $\tau_O = 0$. This defines $C_O = 0$, so that Eqn. (89) turns into

$$\tau_O = \left(1 - \frac{M_\otimes}{R_O} - \frac{V_{O\perp}^2}{2}\right)t \qquad (90)$$

Inverting it we obtain

$$t = \left(1 + \frac{M_\otimes}{R_O} + \frac{V_{O\perp}^2}{2}\right)\tau_O \qquad (91)$$

Substitution of (91) into (88) produces the relation between τ_S and τ_O

$$\begin{aligned}
\tau_S = &\left(1 + \frac{M_\otimes}{R_O} + \frac{V_{O\perp}^2}{2} - \frac{3}{2}\frac{M_\otimes}{a}\right)\tau_O \\
&- 2M_\otimes^{\frac{1}{2}} a^{\frac{1}{2}} e\left[\sin u\left(\left(1 + \frac{M_\otimes}{R_O} + \frac{V_{O\perp}^2}{2}\right)\tau_O\right) - \sin u_0\right] + C_S
\end{aligned} \qquad (92)$$

To make $\tau_S = 0$ at $t = 0$, the same procedure as the one described at the end of section IV should be used. But now, taking into account the results of section

VI, the direction of the synchronizing signal does not need to be restricted to the radial one anymore. Such a procedure will allow us to define the constant C_S. As it has been shown in section VI, we can use the approximation of the expression for the retardation

$$\Delta t = \mid \vec{r}_S - \vec{r}_O \mid \tag{93}$$

Let us suppose now that the synchronizing signal was sent by the ground observer at $\tau_O = \vec{r}_O$. To achieve $\tau_S = 0$ at $t = 0$ the satellite clock should be set to

$$\begin{aligned}
\bar{\tau}_S = \left(1 - \frac{3}{2}\frac{M_\oplus}{a}\right) &\left[\left(1 - \frac{M_\oplus}{R_O} - \frac{V_{O\perp}^2}{2}\right)\tau_O + \Delta t\right] \\
- 2M_\oplus^{\frac{1}{2}}a^{\frac{1}{2}}e\sin u &\left[\left(1 - \frac{M_\oplus}{R_O} - \frac{V_{O\perp}^2}{2}\right)\tau_O + \Delta t\right]
\end{aligned} \tag{94}$$

The formulae (88)–(94) provide a complete account of the relation between the rates of clocks of an observer placed at arbitrary latitude on the ground surface and a satellite in a noncircular orbit (in general, with the plane of its orbit inclined with respect to the equatorial plane), and of the procedure of initial synchronization of the clocks.

The formulae contain second order non–accumulating error in the constant of integration (initial synchronization) and fourth order accumulating with time error.

Numerical evaluation of the ranging errors shows that the relativity produced contribution to the errors can be kept within few centimeters for several years.

IX. NULL–STRUT CALCULUS APPROACH.

As it has been mentioned above, our consideration is an idealized one. The main reason for this was to concentrate all of our attention on general the relativistic effects without getting into details that were irrelevant for our present objectives. However, for practical applications these details can be very important. The orbits of the GPS satellites are not precisely circular, the satellites have non–equatorial orbits, etc. In more realistic problems one cannot possibly hope to be always able to find an analytic solution. The way out is, ordinarily, the use of the numerical techniques. But numerical methods usually have an essential drawback — the loss of the geometric interpretation, which in general relativity means the loss of clarity.

As far as we know there is only one exclusion to this rule — the newly developed null–strut calculus[10,11]. The null–strut calculus was developed originally for numerical solution of geometrodynamic problems. We have shown, for the first time, that null–strut calculus can be used also in chronogeometric problems, and, particularly, in the GPS time transfer problems considered above. It gave us a new stimulating boost in further development of the null–strut calculus, both in theory[12] and in applications[13,14,15].

In null–strut calculus spacetime is triangulated by internally flat simplexes (their size depends on the desired resolution). The chronogeometric relations inside of each simplex are these of special relativity. The general relativistic effects are taken care of automatically by self – adaptation of the null–strut simplicial lattice to the geometry of spacetime. In Fig. 9 and Fig. 10, one can see a pictorial representation of the null–strut calculus interpretation of the Doppler shift (cf. section III) and global synchronization (section IV) in a Schwarzschild field. The picture is made in the same space of coordinates as in sections III, IV. The apparent change in size and shape of quasidiamonds consisting of the null–struts (representing the light rays) and vertical timelike struts directed along the timelike Killing vector
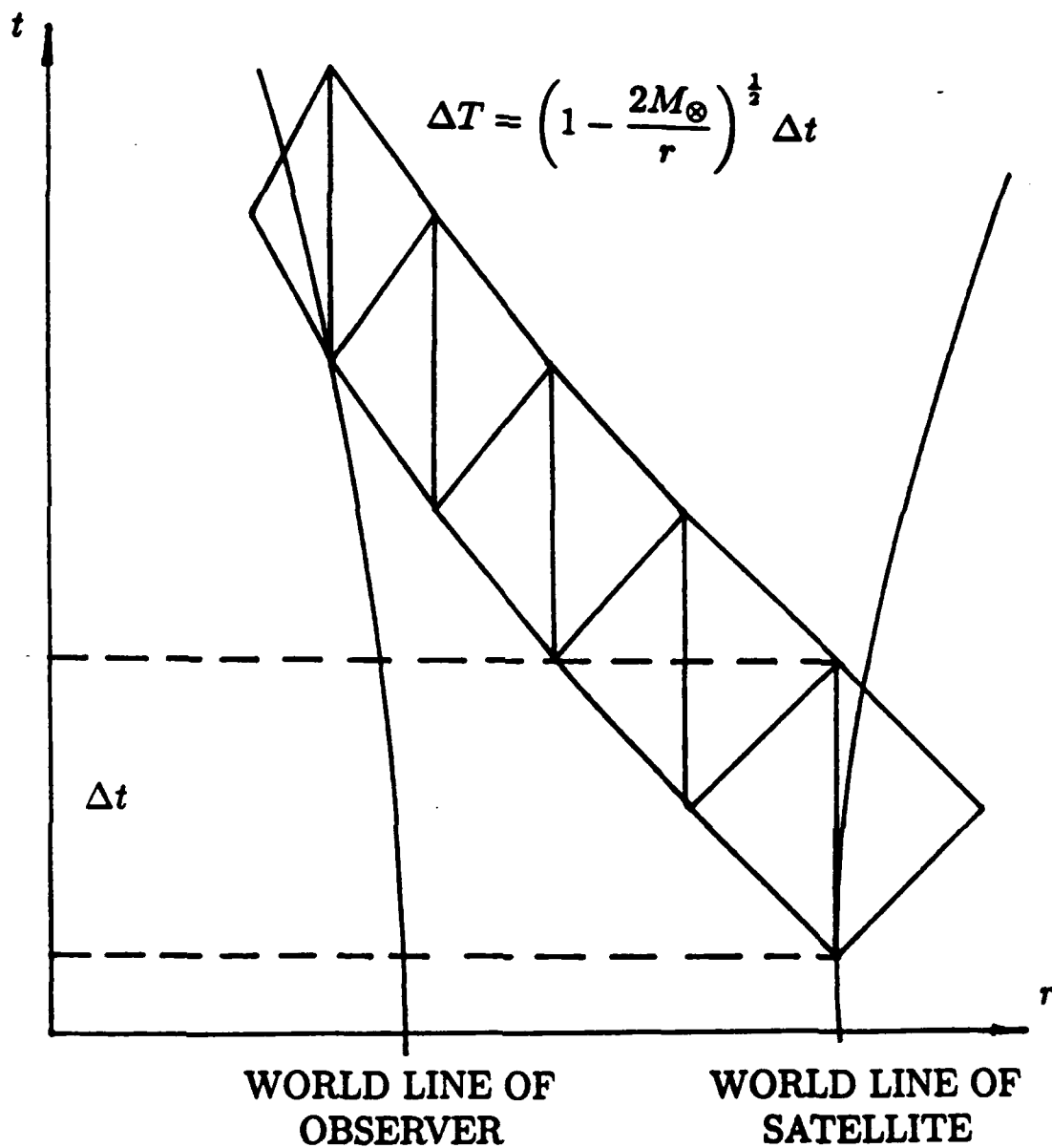
$$\Delta T \approx \left(1 - \frac{2M_\otimes}{r}\right)^{\frac{1}{2}} \Delta t$$

Δt

t

r

WORLD LINE OF
OBSERVER

WORLD LINE OF
SATELLITE

Fig. 9. The null–strut pictorial representation of the Doppler shift calculation in the Schwarzschild field. The vertical dimension Δt of all of the quasidiamonds is the same, but it corresponds to different proper time intervals of clocks resting with respect to the Schwarzschild coordinates $\Delta T = (1 - 2M_\otimes/r)^{\frac{1}{2}} \Delta t$.
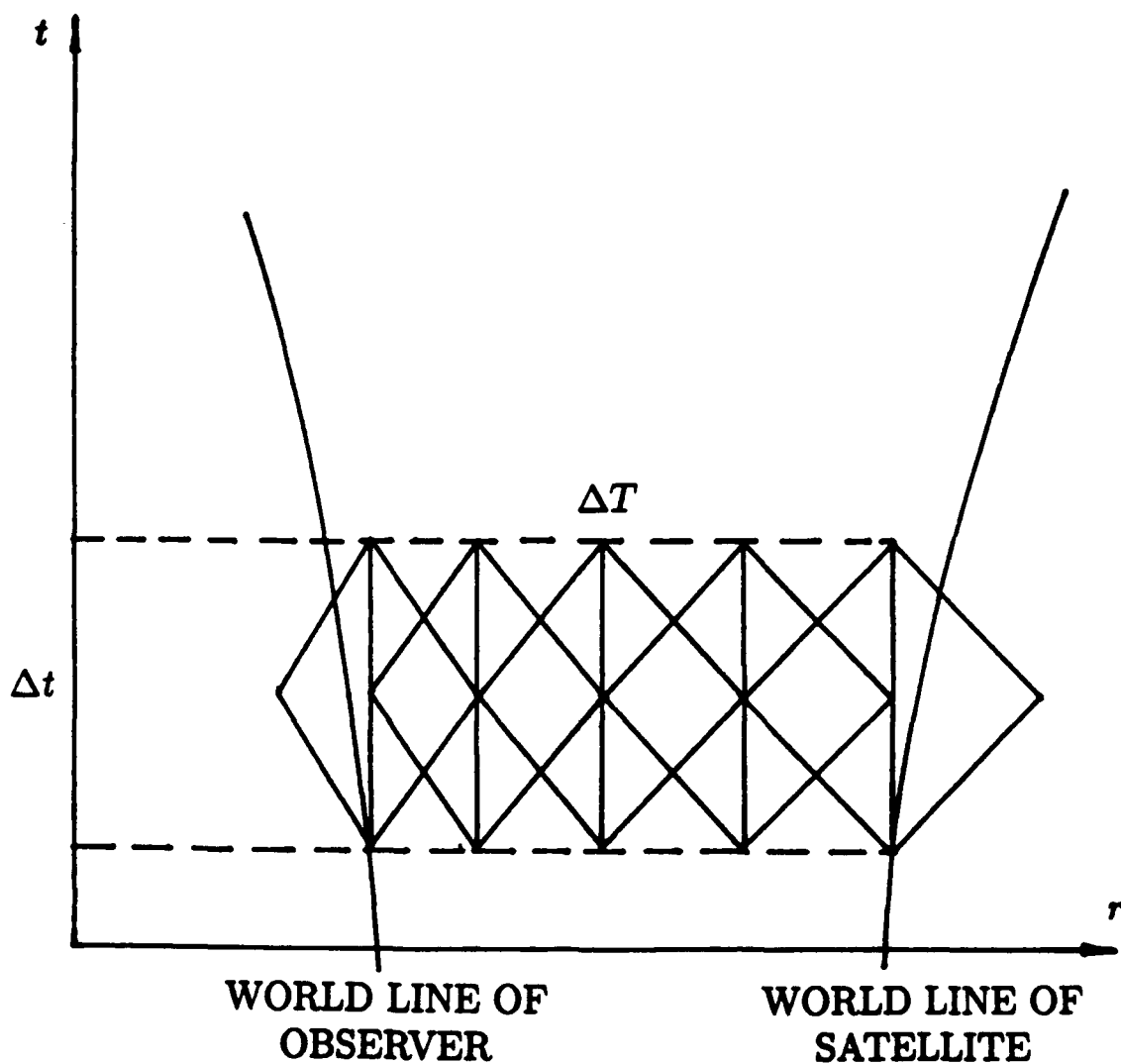
Fig. 10. The null-strut pictorial representation of the relation between the satellite and the ground observer clock rates in the Schwarzschild field. The quasidiamonds of Fig. 9 are dropped down to one level to form an interlocking Schwarzschild simultaneity band.

53-43

field of the Schwarzschild metric is a reflection of the structure of the gravitational field. The quasidiamonds start from the square at spatial infinity and degenerate when approaching the Schwarzschild radius, thus adjusting to the symmetries of the gravitational field.

The results of the null–strut calculations in simple cases coincide with those of sections II, IV[13,14]. But null–strut calculus automatically provides an algorithm to enter into a computer and allows one to obtain numerical solutions in cases which are difficult or impossible to handle analytically. No doubt that in the future development null-strut calculus, as a unique approach having advantages for both numerical techniques and geometry, can become a major tool in the solution and illumination of chronogeometric problems and, in particular, navigational problems.

X. RECOMMENDATIONS.

a. The general relativistic analysis of the Doppler shift and the global synchronization problems in the Schwarzschild field demonstrates clearly the following important features:

(1) The only experimentally measured parameters are the frequencies and the readings of the ground observer clocks and the satellite clocks (proper time);

(2) These parameters are determined at the events (points) of 4-dimensional space-time, and not at points of the 3-dimensional proper space of a reference frame;

(3) Anything else depends on the point of view and the suggested synchronization scheme. It is a matter of interpretation and should be treated as such;

(4) For the Schwarzschild model of the earth's gravitational field, more than one global synchronization scheme, different in the degree of complexity, can be suggested. The synchronization scheme of section IV is much simpler than the one based on the Doppler shift in section III. This simplicity is achieved by the maximal utilization of the Schwarzschild metric symmetries and by the elimination of the terms in the Doppler shift caused by time delay. The scheme is the result of careful analysis of the physical origin of each term in the Doppler shift expression;

(5) Not every interpretation leads to a global synchronization scheme. For instance, an attempt to interpret all the observations in the frame of reference of the ground observer[6] for the purpose of global synchronization is wrong. The world line of the ground observer is curved in such a way that the proper spaces of its frame, corresponding to different moments of his proper time, intersect each other, and consequently cannot be used for global synchronization, even in the special relativistic limit. It is recommended to formulate the problems in four dimensions to avoid mistakes of this kind. In relativity only events are real, and not 3-dimensional positions;

(6) The analysis of the relativistic corrections in the ranging procedure performed in this report shows that the relativistic corrections to the rates of clocks used in modern GPS lead to the fourth order errors in ranging accumulating with time and that the commonly used procedure of initial synchronization (determining the constants of integration) causes an error of second order that does not accumulate. The standard special relativistic time – to – range conversion formula also produces the second order error. This estimate has been completed for the first time in the present report. It implies that the relativistic contribution to the GPS ranging errors can be kept within few centimeters (1–10 cm) for 2–3 years without the resynchronization of the GPS satellite clocks. The analysis was done for an arbitrary latitude of the ground observer and a non–circular orbit of the satellite with an arbitrary orbital plane inclination;

(7) The problem of cross linking based on the satellite – to – satellite ranging can be solved using the techniques of the sections IV–VIII of the present report. The problem is not any harder (in any respect) than the observer – satellite problem. The cross linking, as it follows from the analysis of the sections IV–VIII, should introduce considerable improvements. It will allow one to increase the time intervals between the resynchronizations while maintaining the same precision of the ranging.

b. It is recommended to perform the analysis of the relativity produced positioning errors for the users moving at arbitrary accelerations. Such an analysis requires the application of numerical techniques and an extensive use of computers. The benefit of such research will be the opportunity to further increase the time intervals between the GPS satellites and the users clock resynchronizations. Assuming that both the GPS satellites and users will eventually be equipped with accurate enough clocks capable of working properly under accelerations experienced by the spaceships during launch and maneuver, and that they will also be equipped with devices capable of sufficiently accurate measurements of accelerations, both satel-

lites and users will be able to perform the ranging at a given level of precision for longer time without need of a resynchronization.

c. Numerical methods adopted to the needs of spacetime navigation problems should be further investigated and developed. It is recommended to develop and to implement especially the null–strut calculus methods. Both basic and applied research in this area should be encouraged as much as possible.

REFERENCES.

1. W. A. Miller: Private communication.

2. H. Fligel: Private communication/letter.

3. "Performance Analysis Working Group (PAWG)", chaired by B. Winn, Aerospace Corporation, El Segundo, California, January, 1988.

4. N. Ashby: Private communication, presented at the January 1988 PAWG meeting.

5. R. Matzner: Private communication/letter.

6. C. Alley: Private communication/letter.

7. I. Ciufolini, R. E. Eanes, R. A. Matzner, J. C. Ries, and B. D. Tapley: Private communication.

8. C. W. Misner, K. S. Thorne, J. A. Wheeler: "Gravitation", W. H. Freeman and Co., San Francisco, 1971.

9. J. L. Synge:"Relativity: The General Theory", North–Holland Publishing Co., Amsterdam, 1966.

10. W. A. Miller and J. A. Wheeler:"4–Geodesy", Nuovo Cimento, 8 (1985) 418.

11. W. A. Miller:"Geometric Computation: Null–Strut Geometrodynamics and the Inchworm Algorithm", ed. J. Centrella, Cambridge University Press (1985) 256.

12. A. Kheyfets, N. J. LaFave, W. A. Miller: Pseudo – Riemannian Geometry on a Simplicial Lattice and the Extrinsic Curvature", Phys. Rev., D 39 (1989) 1097.

13. A. Kheyfets, N. J. LaFave, W. A. Miller:"From Geodetic Triangulation to Spacetime Geodesy: Second Order GPS Time Transfer Effects", 16th Gravitational Gradiometry Conference, Sponsored by the USAF Geophysics Laboratory, June, 1988.

14. A. Kheyfets, N. J. LaFave, W. A. Miller: "Doppler Shift and Global Time Synchronization in the Schwarzschild field as Viewed through the Null-Strut Calculus", in preparation.

15. A. Kheyfets, W. A. Miller, J. A. Wheeler: "Null-strut Calculus: The First Test", Phys. Rev. Lett., 61 (1988) 2042.

16. L. R. Gibson: "A Derivation of Relativistic Effects in Satellite Tracking", Report NSWC TR 83-55, April, 1983.

17. A. Kheyfets: "Relativistic Effects in GPS Time Transfer", 1988 USAF-UES SFRP Final Report, July, 1988.

Report # 54
760-7MG-047
Prof. Barry McConnell
No Report Submitted

DEPARTMENT OF CIVIL ENGINEERING
COLLEGE OF ENGINEERING AND TECHNOLOGY
OLD DOMINION UNIVERSITY
NORFOLK, VIRGINIA 23529

PARALLEL AND VECTOR PROCESSING FOR
NONLINEAR FINITE ELEMENT ANALYSIS

By

Duc T. Nguyen, Principal Investigator

Final Report
For the period January through December 1990

Prepared for
Universal Energy Systems
4401 Dayton-Xenia Road
Dayton, OH 45432

Under
F49620-88-C-0053/SB5881-0378
Purchase Order No. S-210-10MG-051
Mr. Rodney C. Darrah

Submitted by the
Old Dominion University Research Foundation
P.O. Box 6369
Norfolk, Virginia 23508-0369

January 1991

ACKNOWLEDGEMENT

I. BACKGROUND ON NONLINEAR ANALYSIS

In structural mechanics, a problem is nonlinear if the stiffness matrix or the load vector depends on the displacements. Nonlinearity in structures can be classed as material nonlinearity (associated with changes in material properties, as in plasticity) or as geometric nonlinearity (associated with changes in configuration, as in large deflections of a slender elastic beam). In general, for a time-independent problem symbolized as [K] (D) = (R), in linear analysis both [K] and (R) are regarded as independent of (D), whereas in nonlinear analysis [K] and/or (R) are regarded as functions of (D).

Many physical situations present nonlinearities too large to be ignored. Stress-strain relations may be nonlinear in either a time-dependent or a time-independent way. A change in configuration may cause loads to alter their distribution and magnitude or cause gaps to open or close. Thus, we see that nonlinear effects may vary in type and may be mild or severe.

An analyst must understand the physical problem and must be acquainted with various solution strategies. A single strategy will not always work well and may not work at all for some problems. Several attempts may be needed in order to obtain a satisfactory result.

II. SCOPES AND OBJECTIVES OF THIS RESEARCH

There are various solution methods for solving systems of nonlinear equations. Each method has its own advantages and disadvantages. A single solution method will not always work

well and may not work at all for some problems. In the next section, several solution methods, such as the Newton-Raphson (N-R), Modified Newton-Raphson (mN-R), and BFGS update methods, are discussed in detail.

The N-R method, due to its quadratic rate of convergence, is an attractive method for solving a system of nonlinear equations. There is a major drawback, however, associated with the N-R method: the coefficient (tangent stiffness) matrix and the right-hand side vector of the above linear system of equations need to be solved repeatedly, and hence, the N-R iteration process can be quite expensive.

To alleviate the above difficulties, the mN-R method is often used. In the mN-R method, the original (symmetrical) tangent stiffness matrix can be used throughout the iteration process, thus, one only needs to form and factorize the (original, symmetrical) tangent stiffness once. This mN-R method, however, requires more iterations to converge to the same pre-scribed tolerance.

The BFGS method provides a compromise between the full reformation of the stiffness matrix performed in the full N-R method and the use of a stiffness matrix from a previous configuration as is done in the mN-R method.

The objectives of this research are three-fold:

    i) identify and parallelize/vectorize the major computational components which are common to most promising solution methods for solving nonlinear structural equations.

ii) incorporate the developed parallel/vector components into promising solution methods such as the N-R, mN-R and BFGS algorithms.

iii) solve a class of practical nonlinear structural mechanic systems in order to evaluate the <u>parallel and vector speed</u> of the developed computer code on the Cray-2 supercomputer.

III. <u>SOLUTION METHODS FOR NONLINEAR STRUCTURAL ANALYSIS</u>

A representative time-independent nonlinear problem can be stated as $[K]\{D\} = \{R\}$, where $\{R\}$ is known and $[K]$ is a function of $\{D\}$ that can be computed for a given $\{D\}$. We are required to compute $\{D\}$, for example, to compute the displacement state associated with known loads. In what follows, we introduce some of the available computational methods. For simplicity, a one-dimensional problem is chosen as the principal example.

Consider a nonlinear spring, Figure 3.1. The source of the nonlinearity is unimportant in the present discussion. We imagine that the spring stiffness k is composed of a constant term k_0 and a term k_N that depends on deformation. Displacement u is caused by load P and is given by the equation

$$(k_0 + k_N)u = P \quad \text{where } k_N = f(u) \tag{3.1}$$

We ask for the value of u when P is given. In order to mimic a realistic problem, we assume that k_N is known in terms of u, and therefore that P can be calculated in terms of u, but that an explicit solution for u in terms of P is not avail-

able. Instead, iterative methods are needed to determine u, as follows.
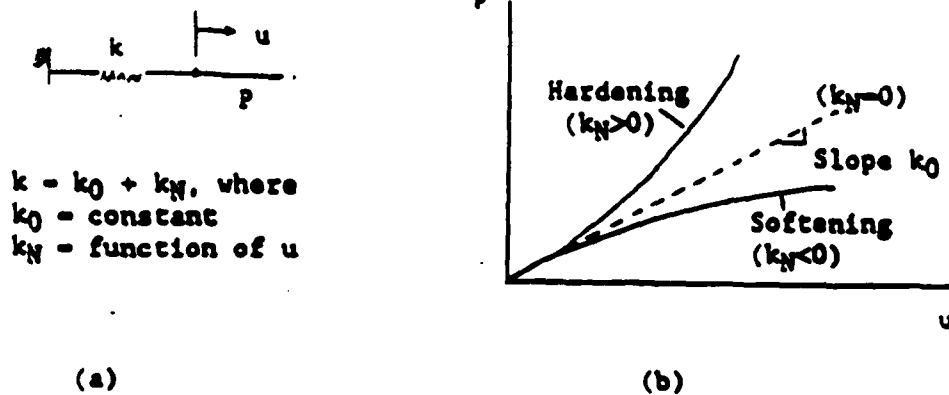

(a)   (b)

Figure 3.1: (a) A nonlinear spring. (b) When u > 0, there is hardening if k_N > 0 and softening if k_N < 0. When u = 0, we assume that k_N = 0.

3.1 Newton-Raphson (N-R) [1]

Imagine that we have applied load P_A and somehow determined the corresponding displacement u_A. That is, from Equation (3.1)

$$(k_0 + k_{NA})u_A = P_A \quad \text{where } k_{NA} = f(u_A) \tag{3.2}$$

The load is now increased to a value of P_B, and the corresponding displacement u_B is sought. A truncated Taylor series expansion of P = f(u) about u_A is

$$f(u_A + \Delta u_1) = f(u_A) + \left(\frac{dP}{du}\right)_A \Delta u_1 \tag{3.3}$$

where

$$\frac{dP}{du} = \frac{d}{du}(k_0 u + k_N u) = k_0 + \frac{d}{du}(k_N u) = k_t \tag{3.4}$$

and k_t is called the tangent stiffness. We seek Δu_1 for which $f(u_A + \Delta u_1) = P_B$. Thus, with $f(u_A) = P_A$ and k_t evaluated at A, Equation (3.3) becomes

$$P_g = P_A + (k_t)_A \, \Delta u_1 \quad \text{or} \quad (k_t)_A \, \Delta u_1 = P_g - P_A \tag{3.5}$$

where $P_g - P_A$ can be interpreted as a load imbalance--that is, as the difference between the applied load P_g and the force $P_A = (k_0 + k_{NA})u_A$ in the spring when its stretch is u_A. The solution process is depicted in Figure 3.2. After computing Δu_1, we update the displacement estimate to $u_1 = u_A + \Delta u_1$. For the next iteration, we obtain a new tangent stiffness $(k_t)_1$ by use of Equation (3.4) with $u = u_1$, and obtain a new load imbalance $P_g - P_1$, where P_1 comes from Equation (3.1) with $u = u_1$. The updated displacement estimate is $u_2 = u_1 + \Delta u_2$, where Δu_2 is obtained by solving $(k_t)_1 \, \Delta u_2 = P_g - P_1$.
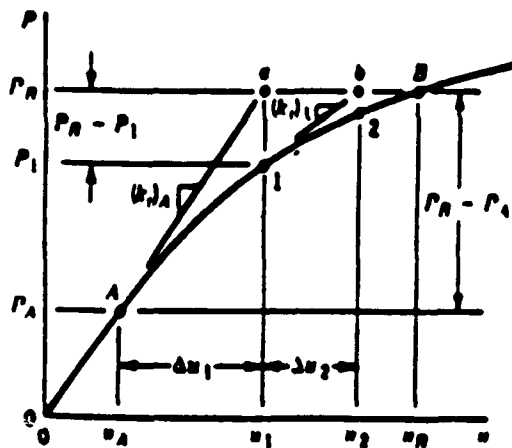

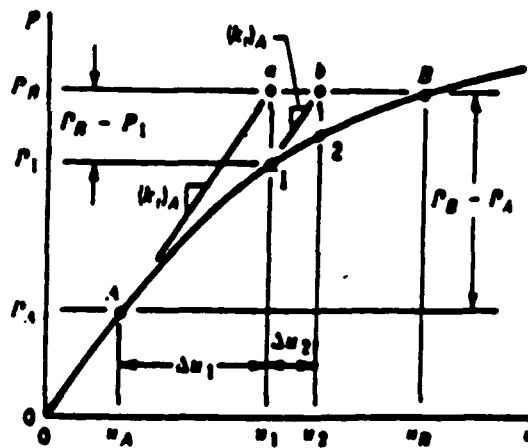**Figure 3.2:** N-R solution for u_B caused by P_B, starting from point A.

Figure 3.3: Modified N-R solution for U_B caused by P_B, starting from point A.

Remarks: Methods discussed in this section extend directly to multiple d.o.f., where $k = k_0 + k_N$ becomes $[K] = [K_0 + K_N]$, P becomes (R), and u becomes (D). In one dimension, if the stiffness can be stated as $k = k_0 + k_N$, the tangent stiffness k_t is

easily obtained in Equation (3.4). Such a simple expression is not available if there are multiple d.o.f. However, in practice, the physics of the problem usually allows us to calculate the tangent-stiffness matrix $[K_t]$. Neither $[K]$ nor $[K_t]$ need to be symmetric in a nonlinear problem, but in some situations symmetry prevails or can be achieved by manipulation.

In a multi-d.o.f. context, N-R iteration involves repeated solution of the equations $[K_t]_i (\Delta D)_{i+1} = (\Delta R)_{i+1}$, where tangent-stiffness matrix $[K_t]$ and load imbalance (ΔR) are updated after each cycle. The solution process seeks to reduce the load imbalance, and consequently (ΔD), to zero.

3.2 Modified Newton-Raphson (mN-R) [1]

This method differs from the N-R method only in that the tangent stiffness either is not updated or is updated infrequently. Thus, in multi-d.o.f. problems, we avoid the expensive repetitions of forming and reducing the tangent-stiffness matrix $[K_t]$. However, more iterative cycles are needed in order to reach a prescribed accuracy. The process is depicted one-dimensionally in Figure 3.3.

3.3 BFGS Update [2]

As an alternative to forms of Newton iteration, a class of methods known as matrix update methods or quasi-Newton methods has been developed for iteration on nonlinear systems of equations. These methods involve updating the coefficient matrix (or rather its inverse) to provide a secant approximation to the

matrix from iteration (i - 1) to i. That is, defining a dis-
placement increment

$$\delta(i) = {}^{t+\Delta t}U(i) - {}^{t+\Delta t}U(i-1) \qquad (3.6)$$

and an increment in the out-of-balance loads

$$\gamma^{(i)} = \Delta R^{(i-1)} - \Delta R^{(i)} \qquad (3.7)$$

the updated matrix ${}^{t+\Delta t}K(i)$ should satisfy the quasi-Newton
equation

$$^{t+\Delta t}K^{(i)} \delta^{(i)} = \gamma^{(i)} \qquad (3.8)$$

These quasi-Newton methods provide a compromise between the full
reformation of the stiffness matrix performed in the full Newton
method and the use of a stiffness matrix from a previous configu-
ration as is done in the modified Newton method. Among the
quasi-Newton methods available, the BFGS (Broyden-Fletcher-
Goldfarb-Shanno) method appears to be most effective, and its use
was first suggested for finite element analysis by Matthies and
Strang.

In the BFGS method, the following procedure is employed in
iteration i to evaluate ${}^{t+\Delta t}U^{(i)}$ and ${}^{t+\Delta t}K^{(i)}$, where ${}^{t+\Delta t}K^{(0)} = {}^{t}K$.

Step 1: Evaluate a displacement vector increment:

$$\Delta U = ({}^{t+\Delta t}K^{-1})^{(i-1)}({}^{t+\Delta t}R - {}^{t+\Delta t}F^{(i-1)}) \qquad (3.9)$$

This displacement vector defines a "direction" for the actual
displacement increment.

Step 2: Perform a line search in the direction ΔU to satisfy "equilibrium" in this direction. In this line search we evaluate the displacement vector,

$$^{t+\Delta t}U^{(i)} = {}^{t+\Delta t}U^{(i-1)} + \beta \Delta U \qquad (3.10)$$

where β is a scalar multiplier, and we calculate the out-of-balance loads corresponding to these displacements ($^{t+\Delta t}R - {}^{t+\Delta t}F^{(i)}$). The parameter β is varied until the component of the out-of-balance loads in the direction ΔU, as defined by the inner product $\Delta U^T({}^{t+\Delta t}R - {}^{t+\Delta t}F^{(i)})$, is small. This condition is satisfied when, for a convergence tolerance STOL, the following equation is satisfied:

$$\Delta U^T({}^{t+\Delta t}R - {}^{t+\Delta t}F^{(i)}) \leq \text{STOL}\ \Delta U^T({}^{t+\Delta t}R - {}^{t+\Delta t}F^{(i-1)}) \qquad (3.11)$$

The final value of β for which (3.11) is satisfied determines $^{t+\Delta t}U^{(i)}$ in (3.10). We can now calculate $\delta^{(i)}$ and $\gamma^{(i)}$ using (3.6) and (3.7) and proceed with the evaluation of the matrix update that satisfies (3.8).

Step 3: Evaluate the correction to the coefficient matrix. In the BFGS method the updated matrix can be expressed in product form:

$$({}^{t+\Delta t}K^{-1})^{(i)} = A^{(i)T}({}^{t+\Delta t}K^{-1})^{(i-1)}A^{(i)} \qquad (3.12)$$

where the matrix $A^{(i)}$ is an (n x n) matrix of the simple form

$$A^{(i)} = I + v^{(i)}w^{(i)T} \qquad (3.13)$$

The vectors $v^{(i)}$ and $w^{(i)}$ are calculated from the known nodal point forces and displacements using

$$v^{(i)} = -\left[\frac{\delta^{(i)^T} \gamma^{(i)}}{\delta^{(i)^T} \,^{t+\Delta t}K^{(i-1)}\delta^{(i)}}\right]^{1/2} \,^{t+\Delta t}K^{(i-1)}\delta^{(i)} - \gamma^{(i)} \qquad (3.14)$$

and

$$w^{(i)} = \frac{\delta^{(i)}}{\delta^{(i)^T} \gamma^{(i)}} \qquad (3.15)$$

The vector $\,^{t+\Delta t}K^{(i-1)}\delta^{(i)}$ in (3.14) is equal to $\beta(\,^{t+\Delta t}R - \,^{t+\Delta t}F^{(i-1)})$ and was already computed.

Since the product defined in (3.12) is positive definite and symmetric, to avoid numerically dangerous updates, the condition number $c^{(i)}$ of the updating matrix $A^{(i)}$ is calculated:

$$c^{(i)} = -\left[\frac{\delta^{(i)^T} \gamma^{(i)}}{\delta^{(i)^T} \,^{t+\Delta t}K^{(i-1)}\delta^{(i)}}\right]^{1/2} \qquad (3.16)$$

This condition number is then compared with some preset tolerance, say 10^5, and the updating is not performed if the condition number exceeds this tolerance.

Considering the actual computations involved, it should be recognized that using the matrix updates defined above, the calculation of the search direction in (3.9) can be rewritten as

$$\Delta U = (I + w^{(i-1)}v^{(i-1)^T}) \dots (I + w^{(1)}v^{(1)^T}) \, \tau K^{-1}(I + v^{(1)}w^{(1)^T}) \qquad (3.17)$$
$$\dots (I + v^{(i-1)}w^{(i-1)^T}) \, [\,^{t+\Delta t}R - \,^{t+\Delta t}F^{(i-1)}]$$

Hence the search direction can be computed without explicitly calculating the updated matrices, or performing any additional costly matrix factorizations as required in the full Newton-Raphson method.

The above BFGS method has already been used effectively in the solution of a variety of nonlinear problems, but the technique constitutes a relatively recent development and significant further experiences in the use of quasi-Newton methods and improvements in their effectiveness can be expected. These methods together with the Newton and other iteration schemes should finally lead to self-adaptive techniques that adjust tolerances and choose solution approaches automatically to obtain an optimum solution of a set of nonlinear finite element equations.

3.4 Convergence Criteria

If an incremental solution strategy based on iterative methods is to be effective, a realistic criteria should be used for the termination of the iteration. At the end of each iteration, the solution obtained should be checked to see whether it has converged within preset tolerances or whether the iteration is diverging. If the convergence tolerances are too loose, inaccurate results are obtained, and if the tolerances are too tight, much computational effort is spent to obtain needless accuracy. Similarly, an ineffective divergence check can terminate the iteration when the solution is not actually diverging or force the iteration to search for an unattainable solution. The

objective in this section is to discuss briefly some convergence criteria.

Since we are seeking the displacement configuration corresponding to time $t+\Delta t$, it is natural to require that the displacements at the end of each iteration be within a certain tolerance of the true displacement solution. Hence, a realistic convergence criterion is

$$\frac{\| \Delta U^{(i)} \|_2}{\| {}^{t+\Delta t}U \|_2} \leq \epsilon_D \qquad (3.18)$$

where ϵ_D is a displacement convergence tolerance. The vector ${}^{t+\Delta t}U$ is not known and must be approximated. In the solution of some problems, it is appropriate to use in (3.18) the last calculated value ${}^{t+\Delta t}U^{(i)}$ as an approximation to ${}^{t+\Delta t}U$. However, in other analyses the actual solution may still be far from the value obtained when convergence is measured using (3.18) with ${}^{t+\Delta t}U^{(i)}$. This is the case when the calculated displacements change only a little in each iteration, but continue to change for many iterations.

A second convergence criterion is obtained by measuring the out-of-balance load vector. For example, we may require that the norm of the out-of-balance load vector be within a present tolerance, ϵ_f of the original load increment

$$\| {}^{t+\Delta t}R - {}^{t+\Delta t}F^{(i)} \|_2 \leq \epsilon_f \| {}^{t+\Delta t}R - {}^{t}F \|_2 \qquad (3.19)$$

3.5 <u>Parallel and Vector Procedures for Nonlinear Finite</u>
<u>Element Analysis</u>

A closer look to <u>all</u> the above solution methods will indi-
cate the following <u>basic computational steps involved:</u>

Step 1:   Guess the <u>initial</u> solution point, $u = u^{(0)}$

Step 2:   Compute the tangent stiffness matrix $K_T(u)$

Step 3:   Compute the unbalanced, or residual vector

Step 4:   Compute the change in the solution, Δu, by solving
the associated system of linear equations

Step 5:   Compute the new, improved solution

$$u^{(i+1)} = u^{(i)} + \Delta u^{(i)}$$

Step 6:   Convergence check.

If YES, then STOP.

If NO, then Return to step 2 (for N-R method or
BFGS method)

or Step 3 (for mN-R method)

The computation involved in Steps 1 and 6 is so trivial and
insignificant. For this reason, only Steps 2-5 of the above
general procedures need to be implemented in a parallel-vector
computer environment.

IV.   STRUCTURAL APPLICATIONS

The parallel-vector versions of Newton-Raphson (N-R),
modified Newton Raphson (mN-R) and BFGS algorithms for nonlinear
structural analysis have been developed and coded on the Cray-2

(or Voyager) supercomputer, using a parallel FORTRAN language FORCE [3]. Two structural examples were used to evaluate the numerical performance of various parallel-vector algorithms. In both examples considered in this section, elapsed (or wall-clock) time in a multiuser (non-dedicated) computer environment for generation and assembly of the structured stiffness matrix, factorization of stiffness matrix, forward and backward substitution, calculation of the unbalanced loads and the total time for calling a particular nonlinear algorithm are reported. The convergence tolerance of 10^{-6} is used for all nonlinear algorithms. In order to save the computational time for the mN-R method, one updates and factorizes the structural stiffness matrix twice (instead of just once!) at the beginning of each load step. Thus, better speed-up factors in all examples can be expected in a truly dedicated computer environment.

4.1 Example 1: A 15 Story x 30 Bay Three-Dimensional
       Truss Structure

A pattern of a three-dimensional truss structure is shown in Figure 4.1. The structure has 5986 truss elements, 992 nodes with 2790 active degree-of-freedom (d.o.f.). The maximum half-bandwidth of the structural stiffness matrix is 202. Young modulus and the cross-sectional area for each member are 19 * 10^6 and 2, respectively. The elapsed time for this example is presented in Table 4.1.
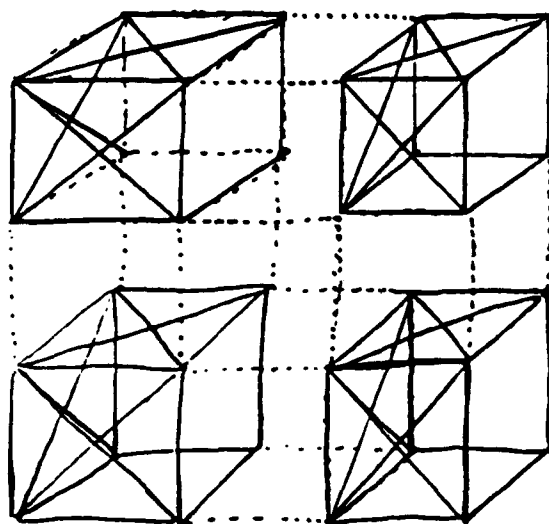
Figure 4.1   A Three-Dimensional Truss Structure

4.2 **Example 2:** A 12 Story x 25 Bay Three-Dimensional Truss
    Structure

A pattern of a three-dimensional truss structure is shown in
Figure 4.1. The structure has 4012 truss elements, 676 nodes
with 1872 active d.o.f. The maximum half-bandwidth of the
structural stiffness matrix is 172. Young modulus and the cross-
sectional area for each member are the same as given in the
previous example. The elapsed time for this example is presented
in Table 4.2.

4.3 **Discussion of the Numerical Results**

Based upon the numerical results obtained from the 2 struc-
tural examples, the following observations can be made:

(a) In a sequential computer environment, the solution time
    for generation and assembly of the structural stiffness
    matrix is significantly smaller than the factorization
    time. However, the time ratio of (generation and
    assembly)/(factorization) is relatively large in a
    parallel-vector computer environment (even for the case
    of 1 processor). This is due to the application of
    loop-unrolling technique and parallel processing equa-
    tion solver [4,5].

(b) The time required for generation and assembly of the
    structural stiffness matrix is decreased as the number
    of processors is increased [6].

(c) For linear static analysis, the solution time for the
    forward and backward elimination phase for solving a

system of linear equations is insignificant as compared to the solution time for the factorization phase. For nonlinear static analysis, however, the forward and backward solution time can be a major component of the total nonlinear analysis time. This is especially true if the mN-R is employed.

(d) Good speed-up factors have been achieved in parallel computation of the unbalanced loads.

(e) Despite the small half-bandwidth (and a non-dedicated computer environment) in both examples, the speed-up factors for parallel factorization are still reasonably good.

(f) In a sequential computer environment, the mN-R is usually much more effective than the N-R method. In a parallel-vector computer environment, however, the N-R method is comparable to the mN-R method (refer to the "total time" column in Table 4.2).

(g) In a parallel-vector computer environment, the BFGS method seems to be the slowest one, as compared to the N-R and mN-R methods.

Finally, it should be mentioned here that the mixed Choleski-S.O.R. parallel-vector algorithm has been successfully tested for a class of (artificial) systems of nonlinear equations [7,8]. For nonlinear structural analysis, however, the parallel-vector S.O.R. method is too slow as compared to the Choleski method. For this reason, the elapsed time for solving the resulting

system of linear equations (of the 2 nonlinear structural examples considered in this research) is not reported.

V. CONCLUSION AND FUTURE RESEARCH

Parallel-vector versions for three promising nonlinear structural analysis methods (the N-R, mN-R and BFGS algorithms) have been developed and coded in a parallel FORTRAN language. Based upon the numerical results from two nonlinear structural applications, the following conclusions can be made.

(a) The mN-R method is often preferred over the N-R method in a sequential computer environment. These two methods, however, are comparable in a parallel-vector computer environment. The parallel-vector BFGS method is slower than the N-R and mN-R methods.

(b) Contrary to linear static analysis, the forward and backward solution time can be a major component in nonlinear static analysis. This is especially true if the mN-R is employed.

(c) Contrary to the sequential nonlinear structural analysis algorithms, the solution time (of parallel algorithms) for generation and assembly of the structural stiffness matrix cannot be considered as an insignificant factor as compared to the factorization time.

(d) Significant reduction in elapsed (non-dedicated) time can be achieved for generation/assembly the structural stiffness matrix, factorization, computation of the

unbalanced loads and the total elapsed time for the entire nonlinear analysis subroutine. In a truly dedicated parallel computer environment, further time reduction can be expected as the number of processors is increased.

Based upon the very promising results obtained from this current research, future research topics which need further investigation include:

1.   Development of Mixed (direct) Choleski - (iterative) Conjugate Gradient [9] parallel-vector algorithm for solving the resulting system of linear equations aris- ing from nonlinear structural applications.

2.   Development of parallel-vector forward/backward proce- dures for the Choleski method. In the shared memory type of supercomputers (such as Convex, the IBM-3090, the Cray-2 and the Cray Y-MP), it has been concluded that the forward/backward solution time can be reduced most effectively by using a single processor (with loop-unrolling technique to exploit the vector facili- ty) rather than using multiple processors [5]. This is due to the fact that the overhead cost for parallel synchronization is more significant than the gains offered by parallel processing. To the author's knowl- edge, this research topic (which is crucially important for nonlinear structural analysis, structural dynamics

and structural optimization) remains to be the unsolvable one in the literature.

3. Development of mixed NR-mNR-BFGS-Arc Length parallel vector algorithm for nonlinear structural analysis. There is no absolutely superior algorithm for all cases. For several cases considered in the 2 examples of this research work, the parallel-vector version of the mN-R method seems to be slightly faster than its N-R counterpart. However, there was a case where the N-R is faster than the mN-R method (see the "total time" column in Table 4.2). A mixed NR-mNR-BFGS-Arc Length algorithm can be tailored to improve the computational speed and the robustness for general nonlinear structural applications.

Cray-2 Time (in seconds)

Method	No. of Processors	Generation and Assembly	Factorization	Forward and Backward	Unbalanced Loads	Total	No. of Load Steps	Total No. Iterations
N-R	1	7.289739	11.299010	0.647	0.301538	20.574693	4	16
N-R	2	4.017558	5.922235	N/A	0.157395	11.337364	4	16
N-R	3	2.533429	4.504895	N/A	0.111726	8.308795	4	16
N-R	4	2.079615	4.323314	N/A	0.091662	7.668497	4	16
mN-R	1	3.643919	5.647862	1.900	0.7545543	14.963744	4	47
mN-R	2	2.007550	2.984869	N/A	0.393237	8.923626	4	47
mN-R	3	1.266236	2.255145	N/A	0.279542	7.099620	4	47
mN-R	4	1.040888	2.044850	N/A	0.230829	6.670830	4	47

Table 4.1. Elapsed Time for A 15 Story x 30 Bay Truss Structure

Cray-2 Time (in seconds)

Method	No. of Processors	Generation and Assembly	Factorization	Forward and Backward	Unbalanced Loads	Total	No. of Load Steps	Total No. Iterations
N-R	1	4.872125	5.918882	0.377	0.202140	12.082768	4	16
N-R	2	2.690095	3.133200	N/A	0.105780	6.711660	4	16
N-R	3	1.698667	2.562486	N/A	0.077177	5.094622	4	16
N-R	4	1.446265	3.142055	N/A	0.064894	5.424594	4	16
mN-R	1	2.435319	2.958809	0.637	0.1036083	7.510705	4	27
mN-R	2	1.345232	1.567790	N/A	0.158988	4.364714	4	27
mN-R	3	0.849777	1.291503	N/A	0.115734	3.502926	4	27
mN-R	4	0.773199	1.483305	N/A	0.098849	3.986571	4	27
BFGS	1	1.826590	2.219945	0.142	0.303714	7.099551	2	8
BFGS	2	1.008925	1.173850	N/A	0.185504	4.885229	2	8
BFGS	3	0.637118	0.956255	N/A	0.116409	4.249042	2	8
BFGS	4	0.532177	1.141107	N/A	0.097820	4.364281	2	8
N-R	1	1.457941	2.554395	0.095	0.145664	4.465516	2	8
N-R	2	0.803903	1.302397	N/A	0.074639	2.431329	2	8
N-R	3	0.759927	1.178630	N/A	0.091180	2.354374	2	8
mN-R	1	0.365182	0.640151	0.721	N/A	4.610998	2	60
mN-R	3	0.131215	0.223360	N/A	N/A	3.889403	2	60

Table 4.2. Elapsed Time for a 12 Story x 25 Bay Truss Structure

REFERENCES

1. Robert D. Cook, David S. Malkus, and Michael E. Plesha, Concepts and Applications of Finite Element Analysis, Third Edition, John Wiley & Sons (1989).

2. K. J. Bathe, Finite Element Procedures in Engineering Analysis, Prentice-Hall (1982).

3. H.F. Jordan, M.S. Benten, N.S. Arenstorf, and A.V. Ramanan, "Force User's Manual," Dept. of Electrical and Computer Engineering, University of Colorado, Boulder, CO 80309.

4. O.O. Storaasli, D.T. Nguyen, and T.K. Agarwal, "The Parallel Solution of Large Scale Structural Analysis Problems on Supercomputers," proceedings AIAA/ASME/ASCE/AHS/ACS 30th SDM Conference, Mobile, Alabama (April 3-5,1989). Also appeared in the AIAA Journal, Volume 28, No. 7, pp. 1211-1216 (July 1990).

5. O.O. Storaasli, D.T. Nguyen and T.K. Agarwal, "A Parallel-Vector Algorithm For Rapid Structural Analysis on High-Performance Computers," NASA-TM 102614, NASA LaRC (April 1990).

6. M.A. Baddourah, O.O. Storaasli, E.A. Carmona and D.T. Nguyen, "A Fast Parallel Algorithm For Generation and Assembly of Finite Element Stiffness and Mass Matrices," accepted for presentation at the AIAA/ASME/ASCE/AHS 32nd SDM Conference, Baltimore, Maryland (April 8 - 10, 1991).

7. E.A. Carmona, D.T. Nguyen, M.A. Baddourah, and T.K. Agarwal, "Parallel and Vector Procedures for Systems of Nonlinear Equations," presented at the ASCE Structures Congress, Baltimore, Maryland (April 30 - May 3, 1990).

8. D.T. Nguyen, O.O. Storaasli, E.A. Carmona, M. Al-Nasra, Y. Zhang, M.A. Baddourah and T.K. Agarwal, "Parallel-Vector Computation For Linear Structural Analysis and Nonlinear Unconstrained Optimization Problems," to appear in Computing Systems in Engineering, An International Journal.

9. J.S. Arora, Introduction to Optimum Design, McGraw-Hill, 1989.

APPENDIX MAY BE OBTAINED FROM THE AUTHOR OR FROM UES, INC.

RESONANT SCATTERING OF ELASTIC WAVES BY A RANDOM DISTRIBUTION OF SPHERICAL INCLUSIONS IN A GRANULAR MEDIUM

a Mini Grant Report

Duane R. Sanders

Department of Civil Engineering

Texas A&M University

College Station, Texas 77843

(713)870-5154

Vikram K. Kinra

Department of Aerospace Engineering

Texas A&M University

College Station, Texas 77843

(409)845-1667

Introduction

Wave propagation in a granular medium containing elastic scatterers is a subject of basic scientific interest in such areas of investigation as soil dynamics, seismology, underground explosions, soil liquefaction and powder metallurgy. An understanding of the mechanics of a homogeneous elastic granular medium (i.e., one without inclusions) has developed over the last twenty years with significant contributions from Goodman and Corvin [1], Nunziato and Walsh [2], Oda [3] and Nemat-Nasser and Mehrabadi [4]. In this report we consider a homogenous granular medium in which a random distribution of elastic scatterers has been placed. The introduction of the scatterers into the granular medium changes the wave propagation characteristics of the medium due to the scattering of the elastic waves at the interface of the inhomogeneities. For an elastic medium the scattered energy is taken away from the incident wave. However, this energy is not converted into heat; rather it is redistributed into incoherent scattered waves. Because the amplitude of the incident wave is reduced as it propagates, the wave appears to be attenuated. It is well known that the spatial decay of a waveform depends on the ratio of the wave length to the size of the inhomogeneities, with the scattering cross-section approaching zero as the wave length approaches infinity. For a inelastic medium there is an additional attenuation of the incident wave due to the dissipative mechanisms which convert some of the mechanical energy into heat.

There are a number of theoretical approaches for analyzing wave scattering which have been reported in the literature. These approaches can be broadly classified as single-scattering or

multiple-scattering theories. It is generally recognized that single-scattering theories apply only at low volume fractions, i.e., when the distance between scatterers is large. Some of the single-scattering theories which have been significantly noted in the literature are due to Ying and Truell [5], Sayer and Smith [6], Moon and Mow [7], Gaunaurd and Überall [8] and Wu [9]. Some of the most recognized multiple scattering theories are due to Foldy [10], Waterman [11], Waterman and Truell [12], Mal and Bose [13] and Varadan, Bringi, Varadan and Ma [14]. Only a few of these wave scattering models have been experimentally verified for bonded elastic materials with a random distribution of scatterers. These verifications have been performed by Kinra for a number of scatterer/matrix systems: glass/epoxy[15], lead/epoxy[16], glass/PMMA[17] and steel/epoxy[18] and random particulate composites and graphite/epoxy fiber reinforced composites[14]. So far as we know, there is no experimental results which can be used to determine if the wave scattering theories apply to granular materials with a random distribution of scatterers . In view of the fact that confining pressure and interface friction play a significant role in scattering of waves in granular materials, it not clear whether the foregoing theories will apply to this situation.

Objective

The objective of this research is to determine experimentally the phase velocity and attenuation of a scatterer/granular system as a function of frequency at several concentrations and confining pressures.

Experimental Set-up and Procedures

The experimental approach taken in this research was similar to that used by Kinra [15,16] to determine the phase velocity and attenuation for particulate composites· using a direct-contact through-transmission technique. The through-transmission

technique was used to determine the phase velocity in the scatterer/granular system by measuring the time of transmission through an aluminum-scatterer/granule-aluminum transmission path as shown in Fig. 1. The two aluminum blocks in the transmission path and a pair of clamps were used to apply pressure to the granules. To prevent the granules from escaping laterally an open box made of plexiglas (into which the two aluminum blocks fit) was used to restrain the granular material as pressure was applied as shown in Figs. 1 and 2. With this granular specimen holder the thickness of the granules through which the waves travel could be changed and the pressure applied to the granules could be set by the torque applied to the four bolts of the clamping device.

The equipment used in this direct contact through transmission experiments is shown in Fig. 3. The pulse generator (Tektronix Type PG 501) was used in conjunction with the function generator (Wavetek model 190) to produce a toneburst which contained a number of cycles with a desired center frequency(0.5 and 1.0 MHz). The tone-burst was amplified by a radio frequency power amplifier (E.I.N. A150) for a peak to peak voltage between 40 and 80 volts and then applied to the transmitting transducer to produce a longitudinal P-wave which was launched into the transmission path. The signal received by the receiving transducer was displayed on a sampling oscilloscope (Tek. 7704A equipped with a P7001 Processor, a 7A16 Amplifier, a 7B80 Time Base Unit and a 7B85 Delaying Time Base Unit. In order to accurately measure the time of transmission through the sample, the lower trace of the oscilloscope is triggered after a suitable delay and swept at a very fast rate (typically 100 ns/division) using the Tek. 7B85 Delaying Time Base Unit to trigger the Tek 7B80 Time Base Unit. The transducers were coupled to the aluminum block using castor oil.

Experimental Constraints

It was determined that the granular medium is very attenuative and

the transmission coefficient between the aluminum block and the granules was not very high. Therefore, it was necessary to select small granule thicknesses at the higher frequencies, since the attenuation increases with frequency. This constraint on specimen thickness determined the number of cycles which could be used in the toneburst and still maintain a clean first transmission received signal which was well separated from the first reflected signal. For the granule thicknesses used and the frequencies considered this resulted in toneburst which contained from 4 to 7 cycles. An example of the toneburst used for a specimen of 0.41 in. is shown in Fig. 4.

Selection of Granules and Scatterers

Plexiglas granules approximately 0.001 inch in diameter were used as the granular material. The transducers used had center frequencies of 0.5 and 1.0 MHz which produced wavelengths in the granular material ranging from approximately 0.11 inches (2.9 mm) at the lowest confining pressure to 0.037 inches (0.93 mm) at the highest confining pressure. Glass beads were selected as the scatterers with a diameter of 0.45-.5 mm in order to obtain normalized frequencies (Ω=ka) in the range from 0.5 to 1.5. It was important to span this range in order to investigate the importance of the size of the scatterer relative to the wavelength.

Measurement of Phase Velocity

To measure the phase velocity the following procedure was used. First, with the granular specimen in the holder (and at the selected bolt torque, i.e. confining pressure) the time taken for a specific peak in the toneburst to propagate through the aluminum-scatterer/granule-aluminum path is measured. Second, the time taken by the wave to propagate through an aluminum block of a known thickness was measured for the same specific peak in the toneburst. The wave speed in the granules can then be calculated

as follows:

$$\langle C \rangle = \frac{d_2}{t_1 - t_2} - \frac{d_2 C_{al}}{d_1 + d_3 + d_4}$$

where d_1 is the thickness of aluminum block 1, d_2 is the thickness of the granules, d_3 is the thickness of aluminum block 2, d_4 is the thickness of a reference aluminum block and C_{al} is the phase velocity in aluminum.

Discussion of Results

First we will report the results obtained for just the matrix material, i.e., the plexiglas granules without the scatterers. Figure 5 shows the wavespeed as a function of the torque applied to the bolts. Clearly the wave speed is a strong function of the torque. This confirms that for granular materials the confining pressure is a very important parameter for characterizing the phase velocity. Figure 6 shows the wave speed versus frequency at four different torques. The figure shows that the wave speed at 50, 75 and 100 in-lb is independent of frequency over the limited frequency range (0.5-1.2 MHz) considered. At a torque of 25 in-lb there seems to be more scatter in the experimental results. This is believed to be due to a difference in confining pressure (bolt torque) at the lower two frequencies than at frequencies above .8 MHz. This difference in confining pressure (bolt torque) is a result of having to reestablish the bolt torque when the specimen thickness was changed between the lower frequencies (0.3-0.5 Hz) and the higher frequencies (0.8-1.2 Hz). This same experimental procedure did not result in a significant scatter in the wave speed at 50, 75 and 100 in-lbs of torque.

We now introduce a normalized frequency $\Omega = k_1 a = 2\pi n a / C_1$ where a is the inclusion radius, n is the frequency, and C_1 is the longitudinal wavespeed in the matrix. Let the volume fraction of

the inclusions be denoted by $\tilde{C}$. Figures 7 and 8 shows the wavespeed of the scatterer/granular system, $<C>$, normalized by the average speed of plexiglas granules, C_{avg}, versus normalized frequency for a $\tilde{C}=0.1$ volume fraction at torques of 50 and 100 in-lb, respectively. The results shown in these figures are very interesting since they show the phase velocity going through a minimum at a normalized frequency of 0.55 for a torque of 50 in-lb and 0.42 for a torque of 100 in-lb. It can also be seen from both figures that waves appears to remain constant for normalized frequencies from 0.8-1.2. This type of wave speed versus normalized frequency has been reported by Kinra [18] for particulate composites. The frequency at which the sharp transition in wave speed occurs is called the cut-off frequency.

Figures 9 and 10 show $<C>/C_{avg}$ versus normalized frequency for a $\tilde{C}=0.25$ volume fraction at torques of 75 and 100 in-lb, respectively. These results are very similar to that obtained for the 0.1 volume fraction given in Figs. 5 and 6. From these figures it can be seen that the wave velocity goes through a minimum at a normalized frequency of $\Omega = 0.5$ for a torque of 75 in-lb and $\Omega = 0.3$ at a torque of 100 in-lb. For both of these torques the wave speed appears to remain constant for normalized frequencies above 0.8. These observations are consistent with the earlier results of Kinra [16].

Conclusions and Recommendations

The phase velocity of the longitudinal waves have been measured for a plexiglas granular medium with a random distribution of spherical glass scatterers in the ultrasonic frequency range of 0.4-1.2 MHz and at volume fractions of 0.1 and 0.25. The phase velocity over this range investigated has been shown to be frequency dependent (i.e. dispersive) with the presence of a cut off frequency (sharp transition in phase velocity) similar to that which has been previously reported for random particulate composites by Kinra [18]. So far as we know this is the first

Acknowledgements

REFERENCES

[1] Goodman, M.A. and Cowin, S.C.,"A Continuum Theory for Granular Materials," Arch. Rat. Mech. Anal., Vol. 44, 1972

[2] Nunziato, J.W. and Walsh, E.K., "Small Amplitude Wave Behavior in One-Dimensional Granular Solids," J. Appl. Mech., Vol. 44, pp. 559-564, 1977

[3] Oda, M.,"Significance of Fabric in Granular Mechanics," Proceedings U.S.-Japan Seminar on Continuum Mechanical and Statical approaches in the Mechanics of Granular Materials, ed. Cowin, S.G. and Statake, M.G., aknjutsu Bunken Fukyukai, pp. 7-26, 1978

[4] Nemat-Nasser, s. and Mehrabadi, M.M.,"Stress and Fabric in Granular Masses," Mechanics of Granular Materials:New Models and Constitutive Relations, ed. Jenkins, J.T. and Statake, M., Elsevier Sci. Pub., pp.1-8,1983.

[5] Ying, C.F. and Truell, R., "Scattering of Plane Longitudinal Wave by a Spherical Obstacle in an Isotropically Elastic Solid," Journal of Applied Physics, 27(9), pp. 1086-1097, 1956

[6] Sayers, C.M. and Smith, R.L., "Ultrasonic Velocity and Attenuation in an Epoxy Matrix Containing Lead Inclusions," J.Phys. D: Appl. Phys., Vol 16, pp 1189-1194, 1983

time a cut-off frequency has been reported for a granular medium.

The through transmission experimental technique developed during the course of this investigation has proven to be a reliable method for measuring the phase velocity of granular materials. Some suggestions for improving the accuracy of the measurements are given below.

The attempt to measure the attenuation was not successful due to the fact that with the present experimental equipment it was difficult to reestablish the confining pressure after taking the granular specimen holder apart in order to add granules and scatterers.

Recommendations for Future Research:

1) The procedure of establishing the pressure in the scatterer/granular system by the torque applied to the bolts is not adequate in consistently reproducing the same pressure. Pressure sensors should be used to measure directly the pressure in the scatterer/granular system due to the torque applied to the bolts.

2) Due to the low transmission coefficient and the high attenuation characteristics of the scatterer/granular system, transducers with a lower center frequency (0.1 MHz) should be used so that a larger amplitude waves can be induced in the granular materials.

3) Method of analysis which have been developed by Kinra and Dayal [19] for analyzing thin specimens (when the tone-burst signal cannot be cleanly separated) should prove very useful in determining the wavespeed and attenuation in granular media.

[7] Moon, F.C. and Mow, C.C., "Wave Propagation in a Composite Material Containing Dispersed Rigid Spherical Inclusions," Rand Corporation Report, RM-6139-PR, Rand, Santa Monica, Ca., 1970

[8] Gaunaurd, G.C. and Überall, H., "Resonance Effects and the Ultrasonic Effective Properties of Particulate Composites," J. Acoust. Soc. Am., Vol. 74, No. 1, pp. 305-313, 1983.

[9] Wu, R., "Attenuation of Short Period Seismic Waves Due to Scattering," Geophysical Research Letter, Vol. 9, No. 1, pp. 9-12, 1982

[10] Foldy, L.L., "The Multiple Scattering of Waves," Physics Review, 67(3), pp. 107-119, 1945

[11] Waterman, P.C., "New Formulation of Acoustical Scattering," Journal of the Acoustical Society of America, Vol. 45, pp. 1417-1429, 1969

[12] Waterman, P.C. and Truell, R., "Multiple Scattering of Waves," Journal of Applied Mathematical Physics, 2(14), pp. 512-537, 1961

[13] Bose, S.K. and Mal, A.K., "Longitudinal Shear Waves in a Fibers-Reinforced Composite," International Journal of Solids and Structures, Vol. 9, pp. 1075-1085, 1973

[14] Varadan, V.K., Bringi, V.N., Varadan, V.V. and Ma, Y., "Coherent Attenuation of Acoustical Waves by Pair-Correlated Random Distribution of Scatterers with Uniform and Gaussian Size Distributions,"Journal of the Acoustical Society of America, Vol. 73, No. 6, pp. 1941-1947, 1983

[15] Kinra, V.K. and Anand, A., "Wave Propagation in a Random Particulate Composite a Long and Short Wave Lengths," International Journal of Solids and Structures, Vol. 18, pp. 367-

380, 1982

[16] Kinra, V.K. and Li, P., "Resonant Scattering of Elastic Waves by a Random Distribution of Inclusions,"International Journal of Solids and Structures, Vol. 22, No. 1, pp. 1-11, 1986

[17] Kinra, V.K. and Ker, E. "Effective Elastic Moduli of a Thin-Walled Glass Microsphere/PMMA Composite," Journal of Composite Materials, Vol. 16, pp. 117-138, 1982

[18] Kinra, V.K., "Dispersive Wave Propagation In Particulate Composites," Recent Advances in Composites in the United States and Japan, ASTM, STP 864, ed. J.R. Vinson and M. Taya, pp. 309-325, 1985

[19] Kinra, V.K. and Dayal, V., "A New Technique for Ultrasonic-Nondestructive Evaluation of Thin Specimens," Experimental Mechanics, Vol. 28, No. 3, pp. 288-297, 1988
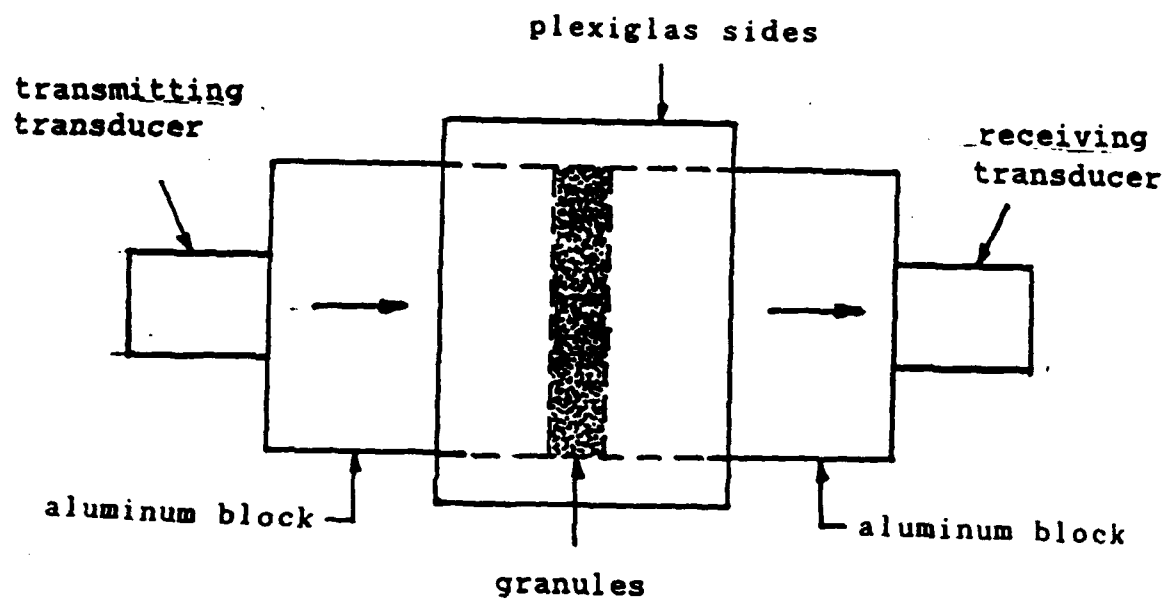
Figure (1) Aluminum-Granule-Aluminum tranmission path
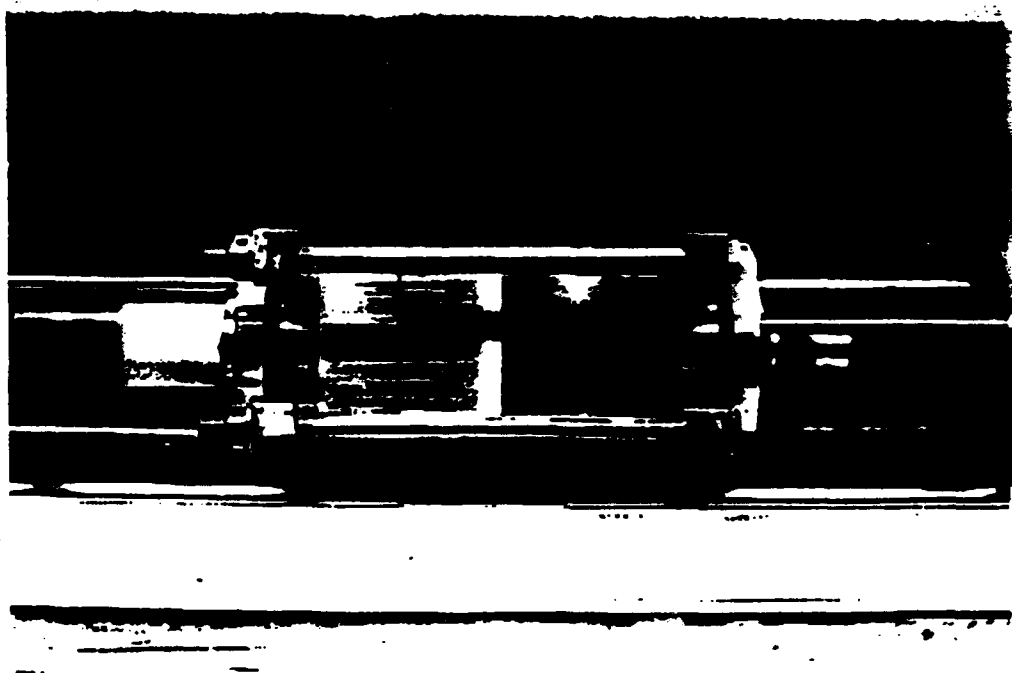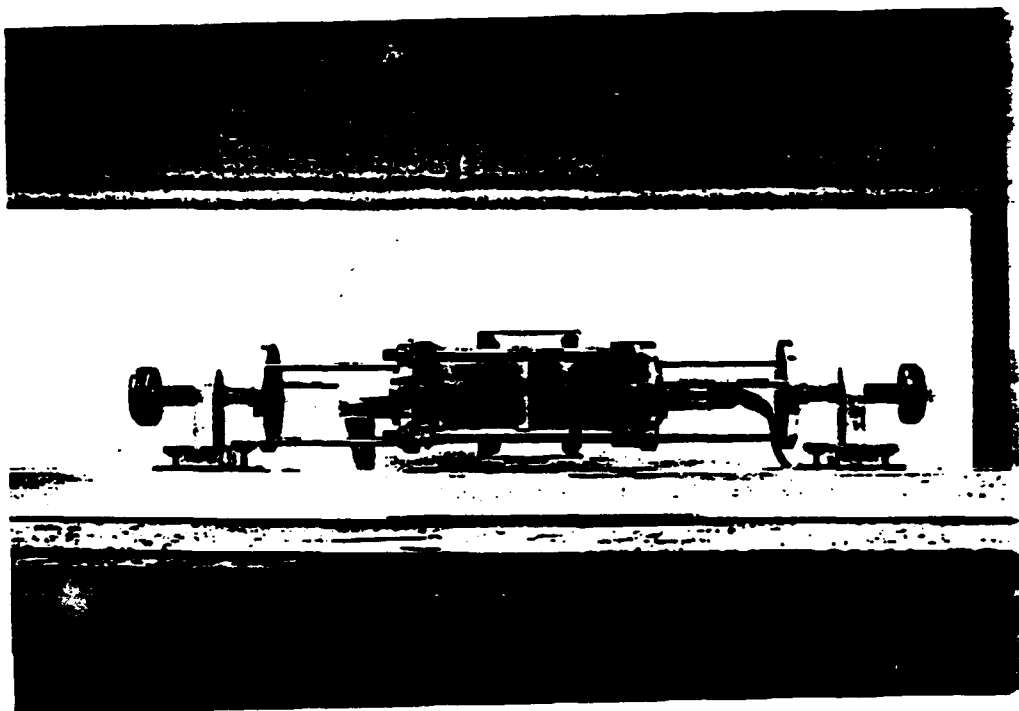
Figure (2) Photographs of scatterer/granule sample holder and transducers
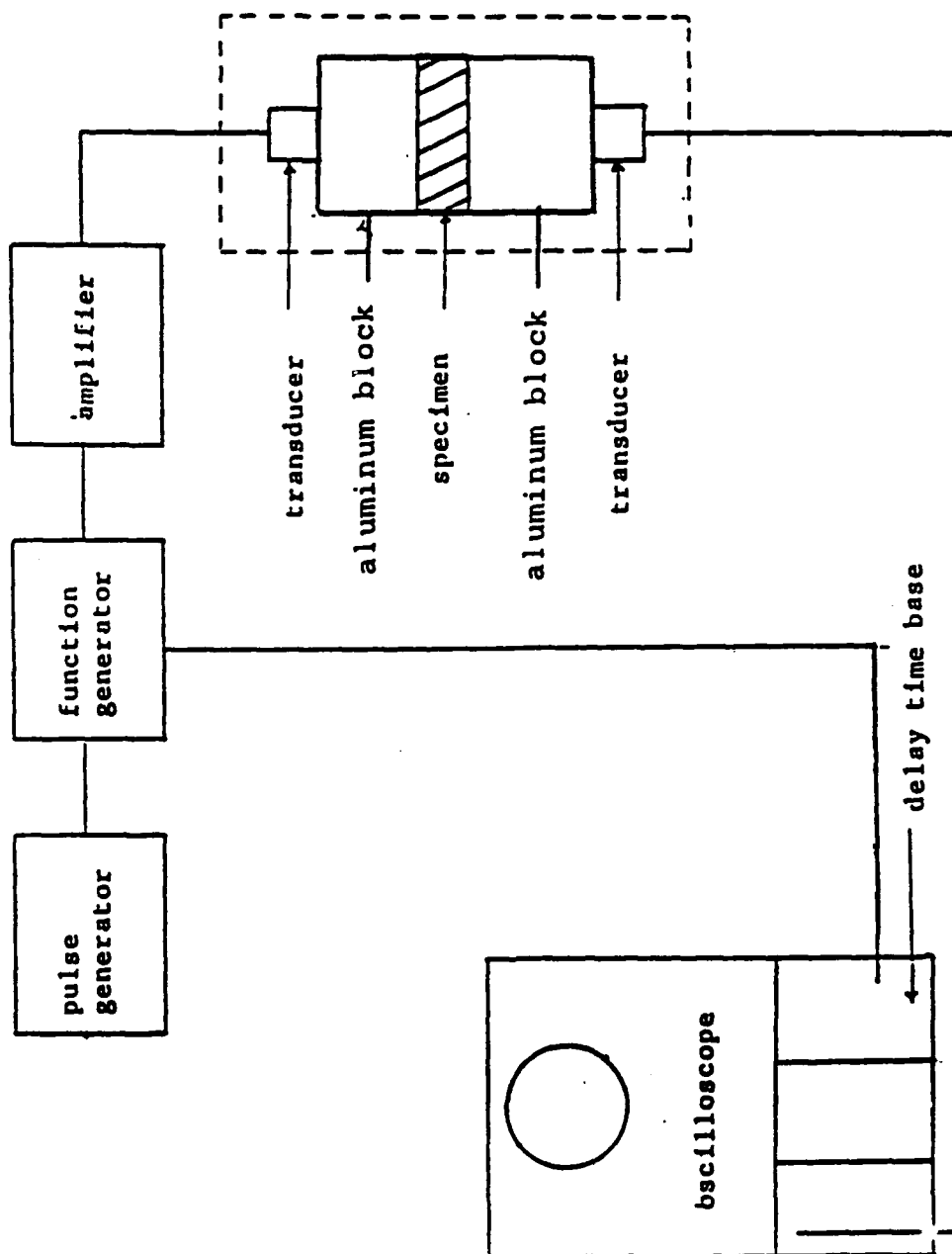
Figure (3) Schematic of experimental apparatus
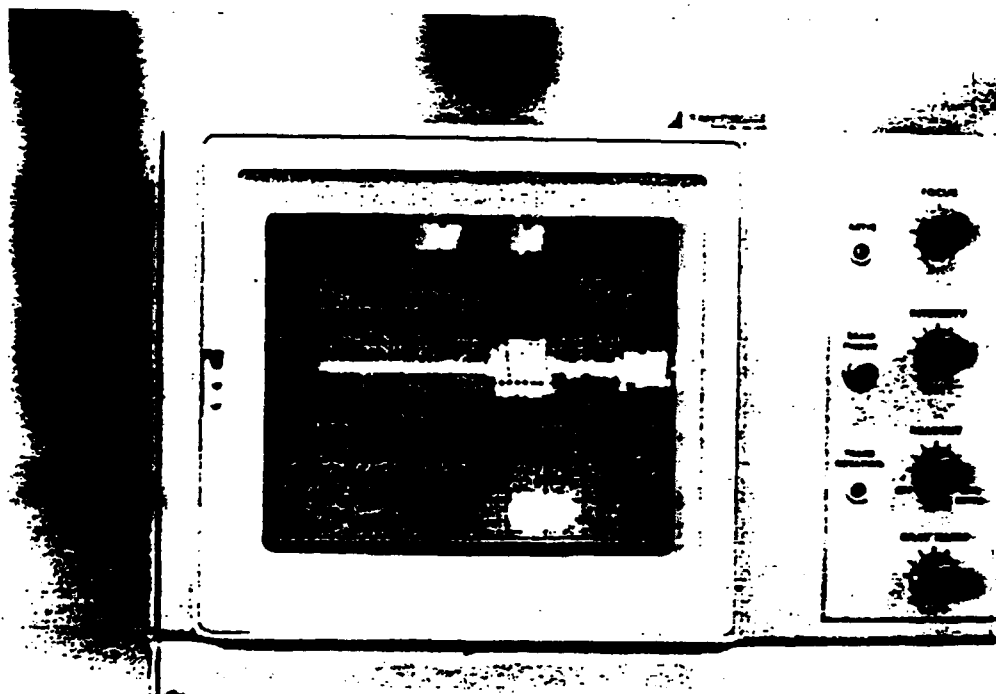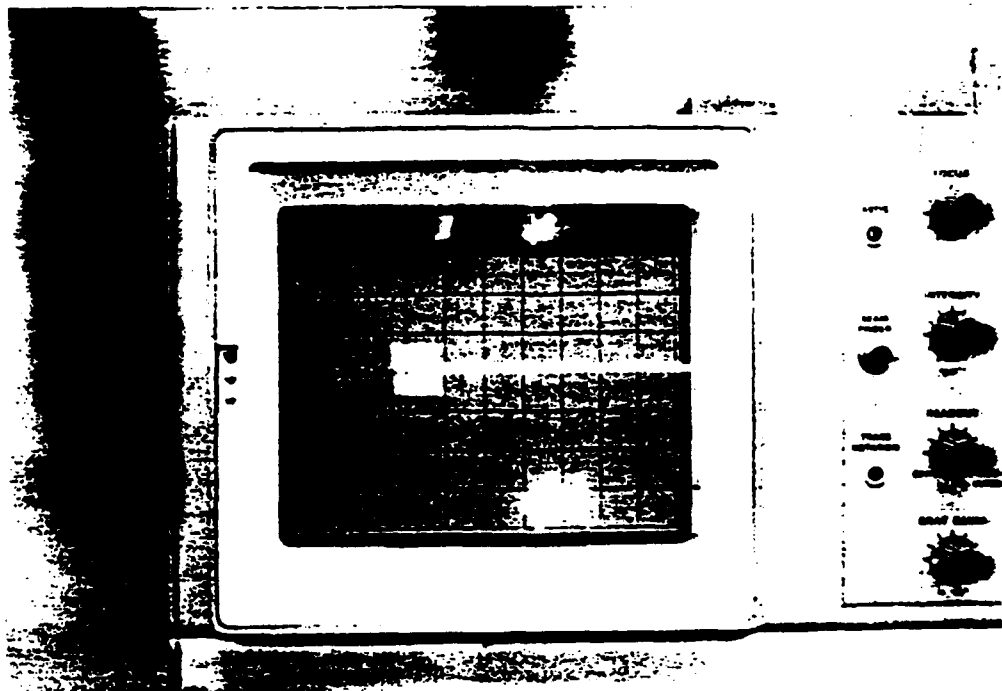
Figure (4) Typical tone burst: a) electrical signal applied to the
transmitting tranducer and b) signal received through
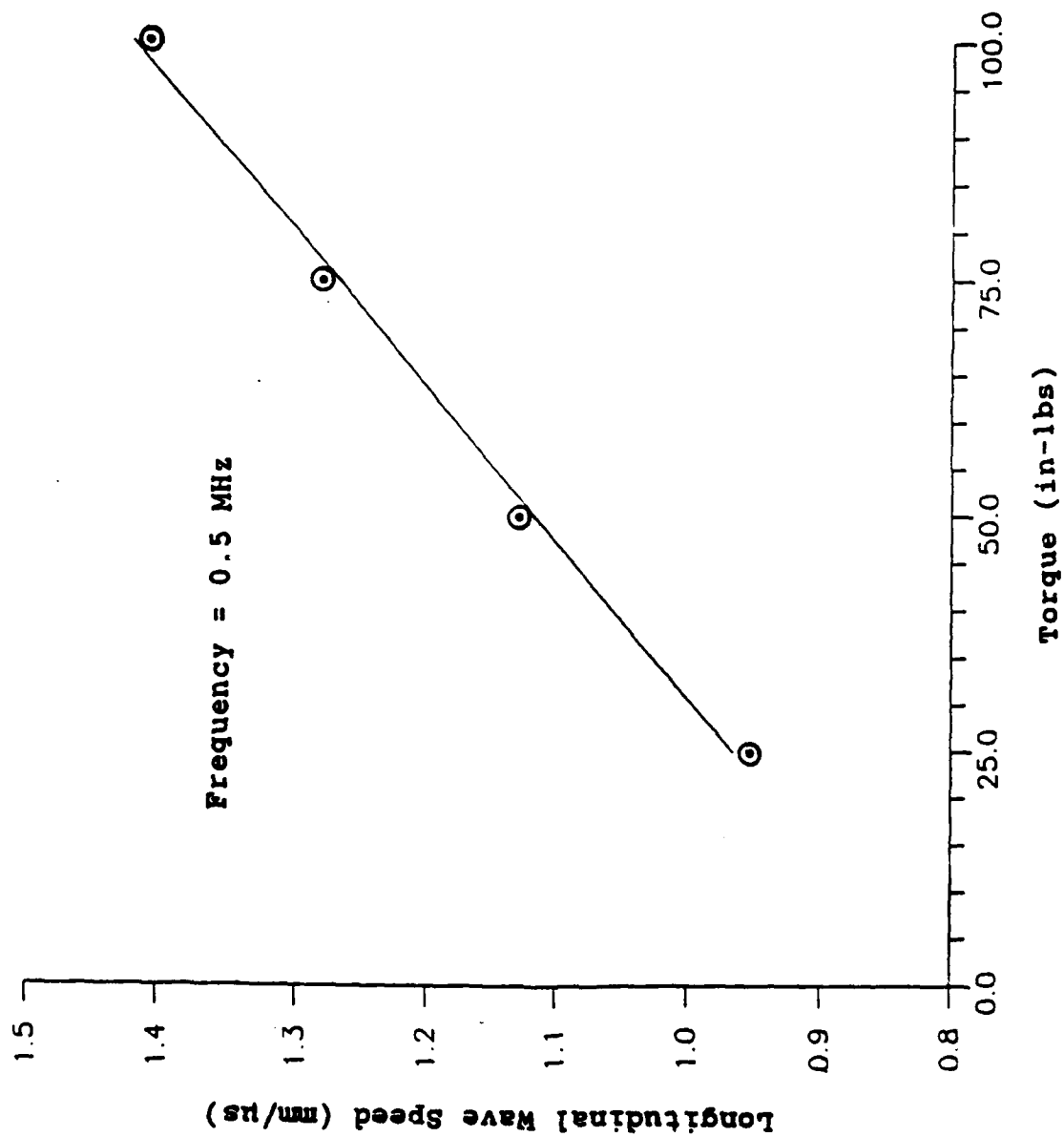the aluminum- grandule-aluminum transmission path

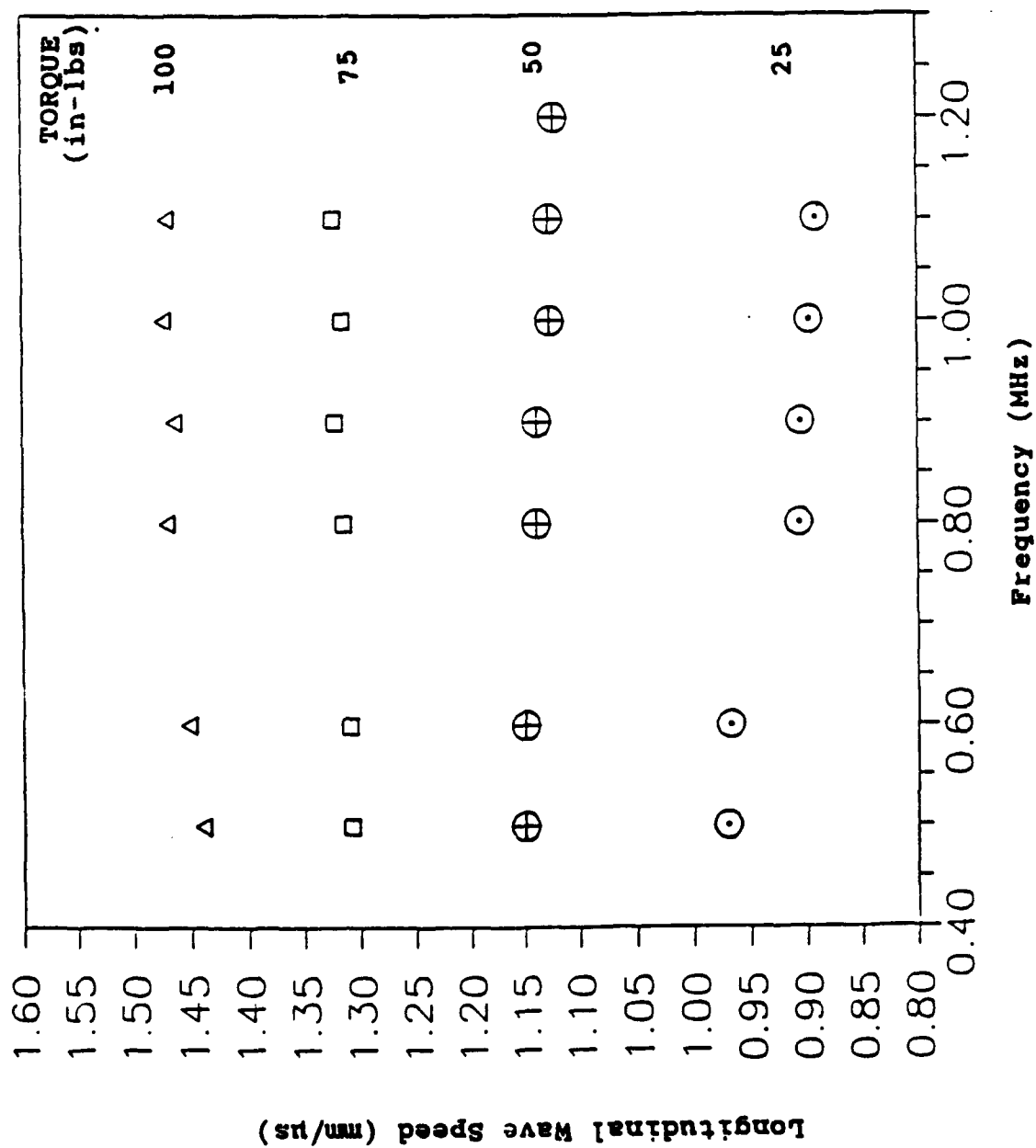**Figure (5) Wave Speed versus torque (granules only)**
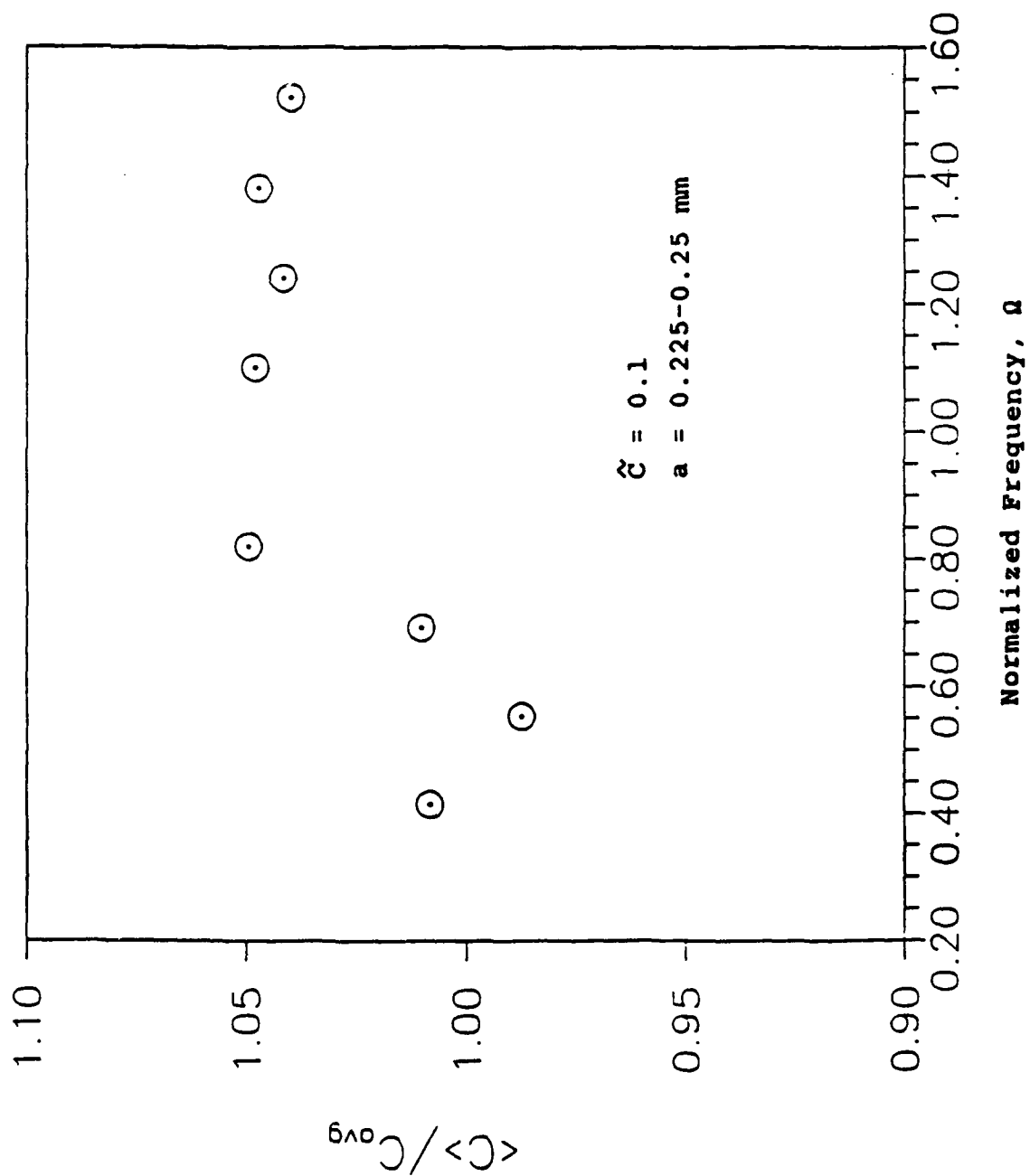
Figure (6) Wave speed versus frequency (granules only)

Figure (7) Wave Speed versus normalized frequency for 0.1 volume fraction (torque = 50 in-lbs)
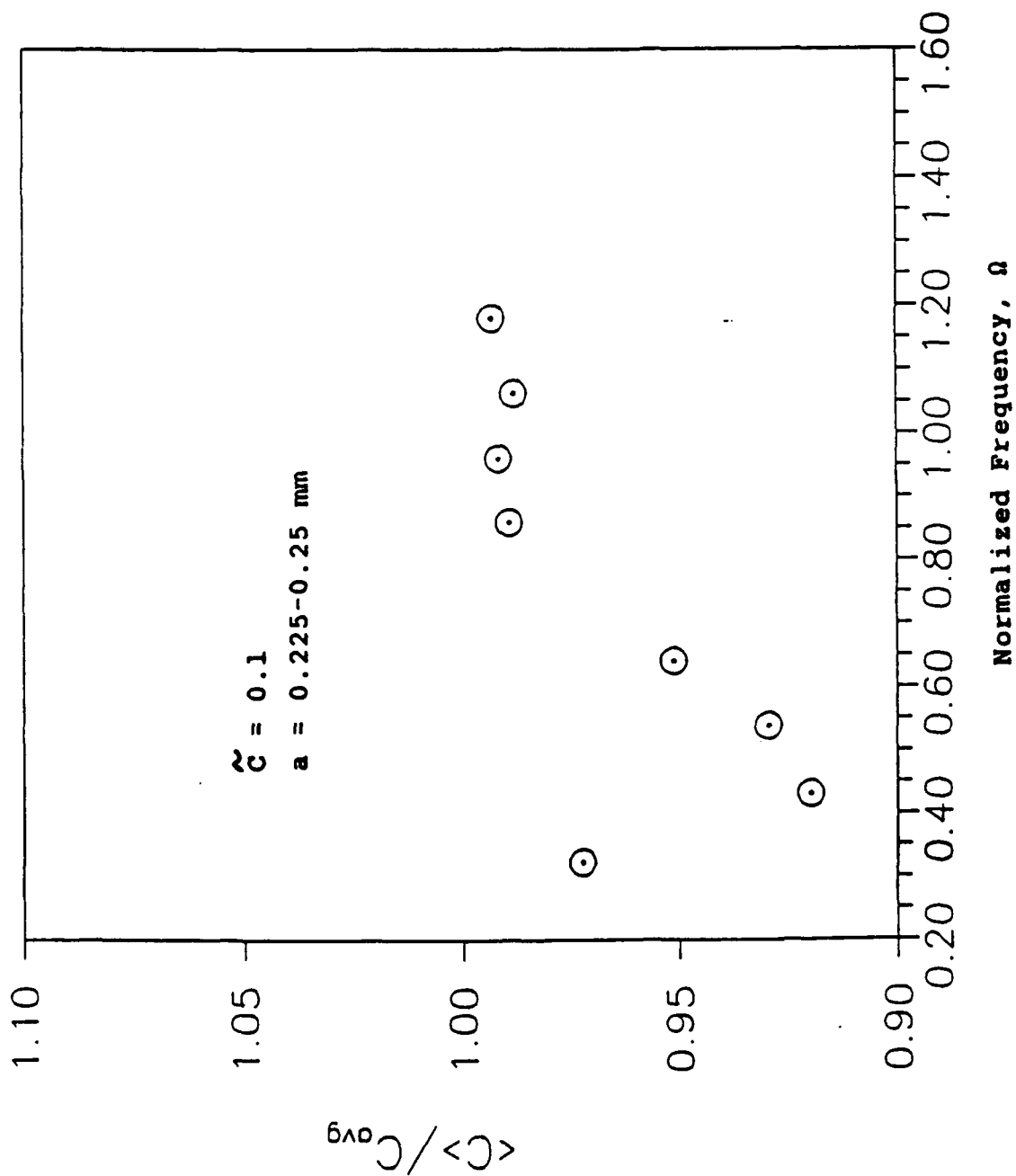
Figure (8) Wave Speed versus normalized frequency
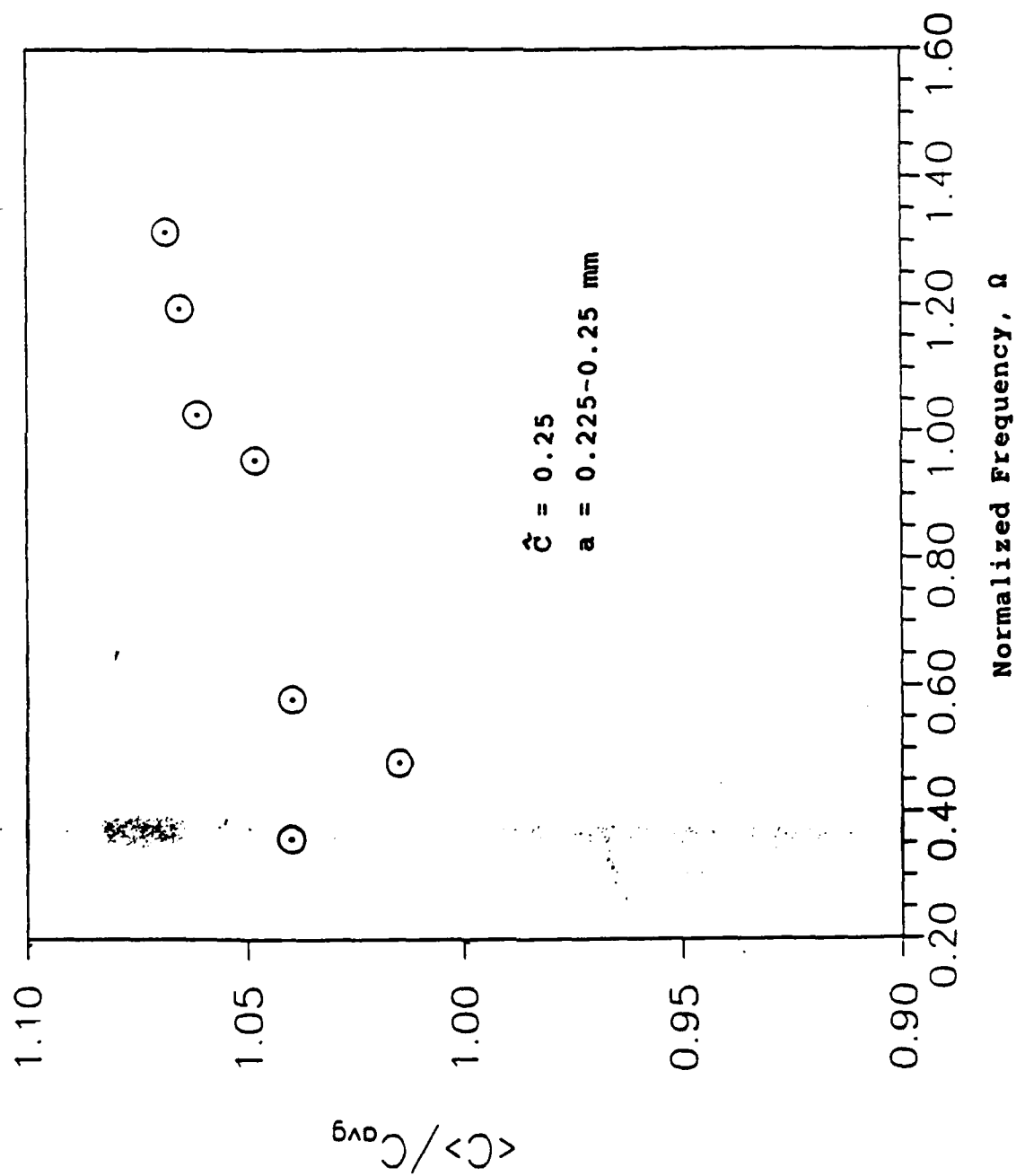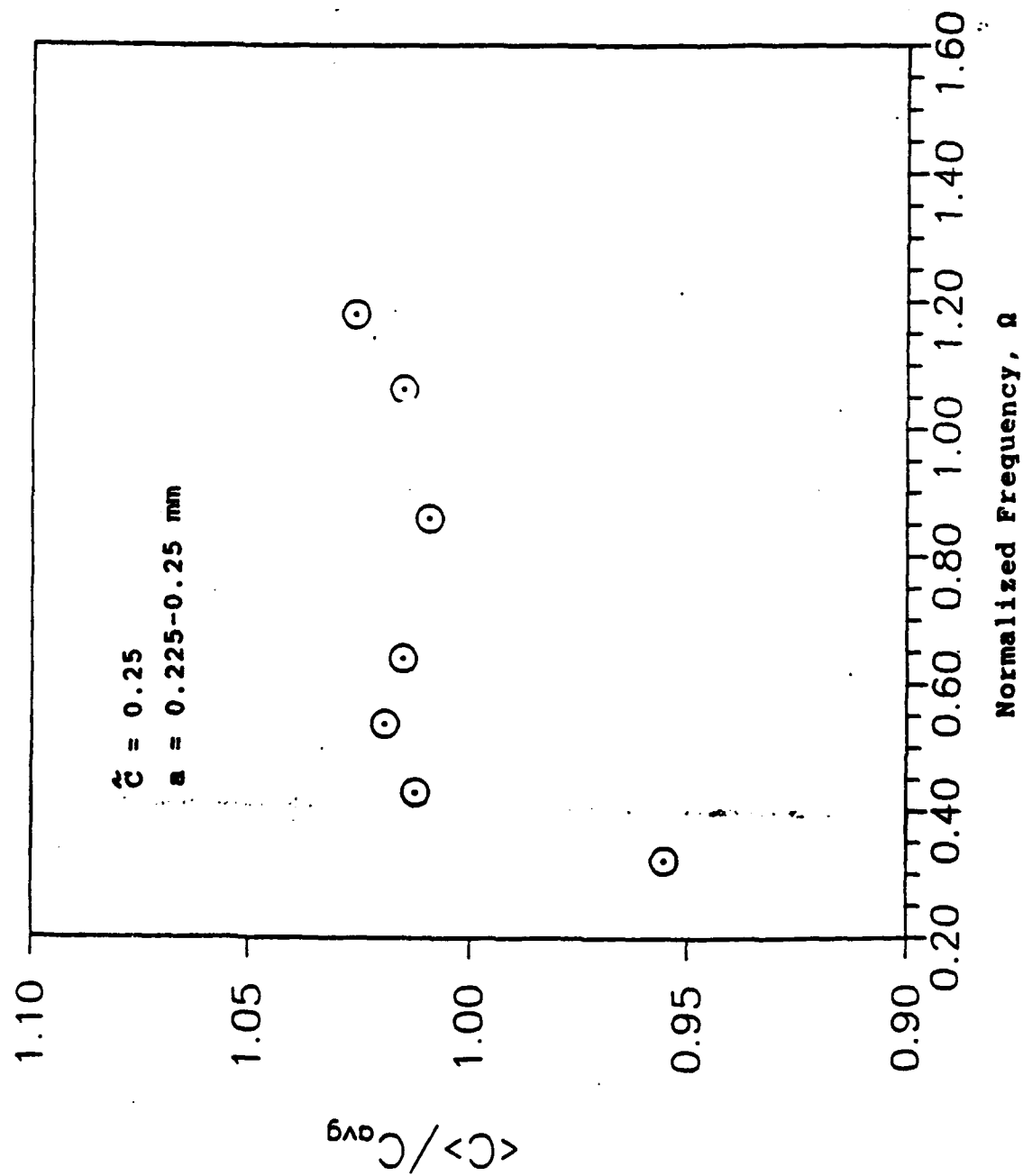for 0.1 volume fraction (torque = 100 in-lbs)

Figure (9) Wave speed versus normalized frequency for 0.25 volume fraction (torque = 75 in-lbs)

Figure(10) Wave Speed versus normalized frequency
for 0.25 volume fraction (torque = 100 in-lbs)